UNIVERSIDAD CARLOS III DE MADRID

ESCUELA POLITÉCNICA SUPERIOR

TESIS DOCTORAL

Dynamic Infrastructure for Federated Identity Management in Open Environments

Autor: Patricia Arias Cabarcos

Director: Dr. Florina Almenares Mendoza

Fecha: 4 de marzo de 2013

TESIS DOCTORAL

Dynamic Infrastructure for Federated Identity Management in Open Environments

Autor: Patricia Arias Cabarcos
Director: Dr. Florina Almenares Mendoza

Firma del Tribunal Calificador:

Firma

Presidente

Vocal

Secretario

Calificación:

Leganés, ___ de _______________ de _____

“If Dynamic federation negotiation and trust management in IdM systems could be achieved it would revolutionize the internet marketplace”

ETSI, 2011

Abstract

Centralized identity management solutions were created to deal with user and data security where the user and the systems they accessed were within the same network or domain of control. Nevertheless, the decentralization brought about by the integration of the Internet into every aspect of life is leading to an increasing separation of the user from the systems requiring access. Identity management has been continually evolving in order to adapt to the changing systems, and thus posing new challenges. In this sense, the challenges associated with cross-domain issues have given rise to a new approach of identity management, called Federated Identity Management (FIM), because it removes the largest barriers for achieving a common understanding.

Due to the importance of the federation paradigm for online identity management, a lot of work has been done so far resulting in a set of standards and specifications. According to them, under the FIM paradigm a person's electronic identity stored across multiple distinct domains can be linked, shared and reused. This concept allows interesting use-cases, such as Single Sign-on (SSO), which allows users to authenticate at a single service and gain access to multiple ones without providing additional information. But also provides means for cross-domain user account provisioning, cross-domain entitlement management and cross-domain user attribute exchange.

However, for the federated exchange of user information to be possible in a secure way, a trust relationship must exist between the separated domains. The establishment of these trust relationships, if addressed in the federation specifications, is based on complex agreements and configurations that are usually manually set up by an administrator. For this reason, the "internet-like" scale of identity federations is still limited. Hence, there is a need to move from static configurations towards more flexible and dynamic federations in which members can join and leave more frequently and trust decisions can be dynamically computed on the fly. In this thesis, we address this issue. The main goal is contributing to improve the trust layer in FIM in order to achieve dynamic federation. And for this purpose, we propose an architecture that extends current federation systems. The architecture is based on two main pillars, namely a reputation-based trust computation module, and a risk assessment module.

In regard to trust, we formalize a model to compute and represent trust as a number, which provides a basis for easy implementation and automation. It captures the features of current FIM systems and introduces new dimensions to add flexibility and richness. The model includes the definition of a trustworthiness metric, detailing the evidences used, and how they are combined to obtain a quantitative value. Basically, authentication information is merged with behavior data, i.e., reputation or history of interactions. In order to include reputation data in the model we contributed with the definition of a generic protocol to exchange reputation information between FIM entities, and its integration with the most widely deployed specification, i.e., Security Assertion Markup Language (SAML).

In regard to risk, we define an assessment model that allow entities to calculate how much risk is involved in transacting with another entity according to its configuration, policies, operation rules, cryptographic algorithms, etc. The methodology employed to define the risk model consists of three steps. Firstly, we design a taxonomy to capture the different aspects of a relationship in FIM that may contribute to risk. Secondly, based on the taxonomy and aiming at developing a computational model, we propose a set of metrics as a basis to quantify risk. Finally, we describe how to combine the metrics into a meaningful risk figure by using the Multiattribute Utility Theory (MAUT) methodology, which has been applied and adapted to define the risk aggregation model.

Furthermore, an also under the MAUT theory, we propose a fuzzy aggregation system to combine trust and risk into a final value that is the basis for dynamic federation decisions.

Formal validation of the above mentioned ideas has been carried out. The risk assessment and decision making are analytically validated ensuring their correct behavior, the reputation protocol included in the trust management proposal is tested through simulations, and the architecture is verified through the development of prototypes. In addition, dissemination activities were performed in projects, journals and conferences.

Summarizing, the contributions here constitute a step towards the realization of dynamic federation, based on the flexibilization of the underlying trust frameworks.

Resumen

Históricamente el diseño de soluciones de gestión de identidad centralizada ha estado orientado a proteger la seguridad de usuarios y datos en entornos en los que tanto los usuarios como los sistemas se encuentran en la misma red o dominio. Sin embargo, la creciente descentralización acaecida al integrar Internet en muchos aspectos de la vida cotidiana está dando lugar a una separación cada vez mayor entre los usuarios y los sistemas a los que acceden. La gestión de identidad ha ido evolucionando para adaptarse a estos cambios, dando lugar a nuevos e interesantes retos.

En este sentido, los retos relacionados con el acceso a diferentes dominios han dado lugar a una nueva aproximación en la gestión de identidad conocida como Federeción de Identidad o Identidad Federada. Debido a la importancia de este paradigma, se ha llevado a cabo un gran trabajo que se refleja en la definición de varios estándares y especificaciones. De acuerdo con estos documentos, bajo el paradigma de identidad federada, la identidad digital de un usuario almacenada en múltiples dominios diferentes puede ser enlazada, compartida y reutilizada. Este concepto hace posibles interesantes casos de uso, tales como el Single Sign-on (SSO), que permite a un usuario autenticarse una sóla vez en un servicio y obtener acceso a múltiples servicios sin necesidad de proporcionar información adicional o repetir el proceso. Pero además, también se proporcionan mecanismos para muchos otros casos, como el intercambio de atributos entre dominios o la creación automática de cuentas a partir de la información proporcionada por otro dominio.

No obstante, para que el intercambio de información personal del usuario entre dominios federados se pueda realizar de forma segura, debe existir una relación de confianza entre dichos dominios. Pero el establecimiento de estas relaciones de confianza, a veces ni siquiera

recogido en las especificaciones, suele estar basado en acuerdos rígidos que requieren gran trabajo de configuración por parte de un administrador. Por esta razón, la escalabilidad de las federaciones de identidad es todavía limitada.

Como puede deducirse, existe una necesidad clara de cambiar los acuerdos estáticos que rigen las federaciones actuales por un modelo más flexible que permita federaciones dinámicas en las que los miembros puedan unirse y marcharse más frecuentemente y las decisiones de confianza sean tomadas dinámicamente *on-the-fly*. Éste es el problema que tratamos en la presente tesis. Nuestro objetivo principal es contribuir a mejorar la capa de confianza en federación de identidad de manera que el establecimiento de relaciones pueda llevarse a cabo de forma dinámica. Para alcanzar este objetivo, proponemos una arquitectura basada en dos pilares fundamentales: un modulo de cómputo de confianza basado en reputación, y un módulo de evaluación de riesgo.

Por un lado, formalizamos un modelo para calcular y representar la confianza como un número, lo cual supone una base para una fácil implementación y automatización. El modelo captura las características de los sistemas de gestión de identidad federada actuales e introduce nuevas dimensiones para dotarlos de una mayor flexibilidad y riqueza expresiva. Se lleva a cabo pues una definición de la métrica de confianza, detallando las evidencias utilizadas y el método para combinarlas en un valor cuantitativo. Básicamente, se fusiona la información de autenticación disponible con datos de comportamiento, es decir, con reputación o historia de transacciones.

Para la inclusión de datos de reputación en el modelo, contribuimos con la definición de un protocolo genérico que permite el intercambio de esta información entre las entidades de un sistema de gestión de identidad federada, que ha sido además integrado en el estándar más conocido y ampliamente desplegado (Security Assertion Markup Language, SAML).

Por otro lado, en lo que se refiere al riesgo, proponemos un modelo que permite a las entidades calcular en cuánto riesgo se incurre al realizar una transacción con otra entidad, teniendo en cuenta su configuración, políticas, reglas de operación, algoritmos criptográficos en uso, etc. La metodología utilizada para definir el modelo de riesgo abarca tres pasos. En primer lugar, diseñamos una taxonomía que captura los distintos aspectos de una relación en el contexto de federación de identidad que puedan afectar al riesgo. En segundo lugar, basándonos en la taxonomía, proponemos un conjunto de métricas que serán la base para cuantificar el riesgo. En tercer y último lugar, describimos cómo combinar

las métricas en una cifra final representativa utilizando el método Multiattribute Utility Theory (MAUT), que ha sido adaptado para definir el proceso de agregación de riesgo.

Además, y también bajo la metodología MAUT, proponemos un sistema de agregación difuso que combina los valores de riesgo y confianza en un valor final que será el utilizado en la toma de decisiones dinámicas sobre si establecer o no una relación de federación.

La validación de todas las ideas mencionadas ha sido llevada a cabo a través del análisis formal, simulaciones, desarrollo e implementación de prototipos y actividades de diseminación.

En resumen, las contribuciones en esta tesis constituyen un paso hacia el establecimiento dinámico de federaciones de identidad, basado en la flexibilización de los modelos de confianza subyacentes.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

"Management of trust, and by extension identity, must be a first order capability."

David D. Clark, 2012

Contents

1.1. Motivation and Objectives

The management of digital identity is said to be located at the core of the Internet economy [OECD Report, 2011], since its evolution can leverage a new wave of innovation explosion. On the Internet nowadays, as well as in other network scenarios, it is likely that each user ends up with multiple credentials and multiple access permissions across different applications provided by different service providers. Our identity data is scattered and these fragmented logins present a challenge in forms of synchronization of shared identities, security, etc. There is a strong need for an intrinsic identity system that is trusted across the web and within enterprises and for unambiguously identifying users.

Recently, **Identity Federation**[1][2] has emerged as a key concept in the identity management field, as a mean of linking a person's electronic identity stored across multiple distinct domains. Thus, the federation model enables users of one domain to securely access resources of another domain seamlessly, without the need for redundant user login processes. Particularly, the most popular use-case in Federated Identity Management (FIM) is Single Sign-On (SSO), which allows users to authenticate at a single site and gain access to multiple ones without providing additional information. Due to the importance of the federation paradigm for online identity management [Chadwick, 2009], a lot of work has been done so far. As a result, the industry and research community have produced a number of standards and specifications [Cantor et al., 2005b] [Goodner and (eds.), 2009] [OpenID, 2007] [Bertocci et al., 2007] [Hammer-Lahav, 2010] representing the fundamental building blocks to accomplish identity federation. However, none of the specifications define a suitable trust model to allow the establishment of dynamic federations.

Trust is a fundamental issue to address scalability. Moreover, the flexibility of every federation framework is tied to the underlying trust model, often poorly defined or even out of the specifications scope. For this reason, new enhanced techniques are required to achieve ad-hoc dynamic federation. Furthermore, the significance of research on this topic has been recently highlighted to the point of stating that *"If dynamic federation negotiation and trust management in IdM systems could be achieved, it would revolutionize the internet marketplace"* [ETSI, 2011]. The work in [Dabrowski and Pacyna, 2008] defines a modular reference architecture that abstracts the different identity layers on the Internet, pointing out the state of implementation. This architecture is depicted in the left side of Figure 1.1. In turn, the right side of the picture shows how existing identity specifications implement particular functionalities. It can be seen that the Trust Layer, which is mandatory for dynamic and ad-hoc inter-federation interactions, is still uncovered.

Motivated by this problem and aiming at contributing to improve the identity landscape, and more specifically the logic trust layer shown in 1.1, we propose "**a dynamic infrastructure for federated identity management in open environments**".

[1] We will use the terms "identity federation" and "federated identity management" indistinctly throughout this dissertation.

[2] Sometimes in the literature, a distinction is made between federated identity and user-centric identity [Hardt, 2005] consider this latter as an evolution of the former. However concepts are not exclusive since federated identity can be implemented empowering user control, thus having federated user-centric systems [Suriadi et al., 2009]. Our proposal applies in any case, because, whether user-centric or not, trust relationships must be always established between the parties sharing identity information.

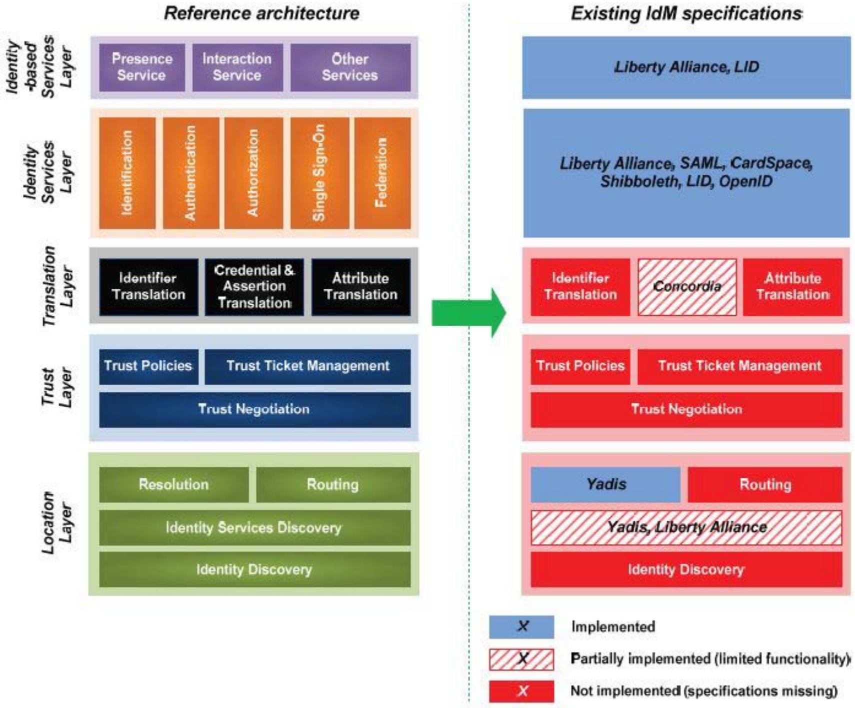


Figure 1.1: Reference framework architecture for Identity Management systems and current state of its implementation (©[Dabrowski and Pacyna, 2008])

On the one hand, with "dynamic infrastructure", we refer to the design of the necessary structures and elements that support the creation of federations in a more agile way, minimizing pre-configuration. What if every time we wanted to send an email, we had to get your IT administrator to coordinate a secure connection between our email server and the email server of the receiving company? Probably if this level of pre-configuration was required, email would not be ubiquitous now. This problem still needs to be solved in FIM to achieve wide scale adoption.

On the other hand, with the term "open environments" we refer to the fact that participation in FIM should be open, market driven, and transparent. Providers may join and leave federations at any moment, they may belong to different domains and may be unknown to each other.

Summarizing, the problem of establishing federations in dynamic and open environments is that current technologies require trust and contractual frameworks to be pre-configured before any interaction between parties takes place. Thus, the initial setup complexity is a high barrier and may not worth adopting these procedures for a short-term collaboration because time and cost will probably not outbalance the rewards of cooperation. Therefore, the main goals of our research are oriented to overcome the limitations of the current static features of FIM systems, and can be summarized in:

- **Minimize dependence on pre-configuration, improving automation.**

  We aim to make entities involved in FIM transactions more autonomous and capable of making decisions to collaborate in a dynamic fashion. The intended contribution towards this goal encompasses the analysis of the gaps and limitations that make current FIM technologies static, and the design of an architecture to cover these limitations. We envision risk and trust evaluation as core aspects of the architecture, which must be considered in decision making. Consequently, the other main goals of this thesis are the development of appropriate risk and trust models for FIM. Having formal models to compute and represent trust and risk as numbers provides a basis for easy implementation and automation.

- **Introduce a risk management model.**

  The introduction of risk analysis enhances security and provides a solid base for deciding whether to cooperate or not with unknown potential partners. It will allow entities to know how much risk is involved in transacting with an entity according to its configuration, policies, operation rules, cryptographic algorithms, etc. Evaluation shall be made on the fly, on a per transaction basis. The contribution shall include the design of metrics to quantify all the risk aspects involved in FIM scenarios, as well as an aggregation model that leads to a meaningful numerical value upon which decisions can be made. This is the most challenging goal, since there are no proposals on risk calculation in FIM.

- **Enrich trust mechanisms.**

  Current FIM frameworks are based on binary decisions (certificate-based trust). We aim to enrich the trust establishment procedures by taking advantage of common knowledge (reputation) and monitoring the evolution of the relationship through time (history of interactions). The contribution shall model the features of current FIM systems and introduce the new mentioned dimensions to add flexibility and richness. The model shall detail the evidences used, and define how to combine them to obtain a final quantitative value.

Furthermore, as it is also an important part in the development of a doctoral thesis, we also aim to achieve the following goals:

- Evaluation and validation of the contributions.
- Dissemination of the results through publication, collaboration in research projects and participation in conferences.
- Identification of new lines of research that can be derived from this work.
- Completion of the writing and public defense of the thesis dissertation.

1.2. Development Plan

With the aim to achieve the goals presented in the previous section, we will follow these steps:

- Gather the bibliography related to FIM in order to study and analyze the existing gaps in regard to trust establishment.
- Design an architecture with the necessary elements to permit dynamic federations based on trust and risk.
- Develop the quantitative trust and risk models, which constitute the main pillars of the architecture.
- Design a decision system that combines the trust and risk values.
- Perform evaluation and validation tests of the proposed architecture and designed modules in order to demonstrate the benefits of the proposal and its feasibility.
- Obtain the main conclusions from the performed research work and identify new research lines to be followed.
- Write and publish papers with the partial results that are obtained during the different phases of the research.
- Write the dissertation document.

1.3. Interest of the Research

With respect to publication and dissemination, the content of this thesis was developed as a research line in two national R&D projects: "España Virtual"[3] and CONSEQUENCE[4]. Both projects included an specific working package for "*Security and Identity Management*", where our ideas on dynamic federation were contributed. In addition, during a research stay at NEC Laboratories Europe, the work on the integration of reputation in FIM was included in the deliverables of an internal business project centered in IdM.

Furthermore, dissemination was also achieved through publication of scientific papers. The main papers that support the interest of the research presented in this thesis are detailed below. For each contribution, we briefly explain the kind and date of publication (i.e., whether conference or journal) and its contents, showing which part of the dissertation they support. It is to note that all the journal papers correspond to journals indexed in the JCR. We also reference other complementary works we have published that, though they do not deal with core aspects of this thesis, are derived from the ideas presented here (e.g., use-cases, application scenarios, etc.). The criterion for ordering the results is their relevance to the dissertation, so more relevant papers are listed first.

Main Contributions to Journals:

1. Title: *A metric-based approach to assess risk for "On Cloud" Federated Identity Management.*
   Authors: P. Arias-Cabarcos, F. Almenares, A. Marín, D. Díaz-Sánchez, R. Sánchez.
   Journal: Springer's Journal of Network and Systems Management, Special Issue on Cloud Computing, Networking, and Service (CCNS) Management. September 2012. (Impact Factor 2011 0.452) [Arias et al., 2012b].

   In this paper we analyze the FIM process and propose a comprehensive taxonomy (starting form the structure outlined in [Arias, 2011]) that helps in the classification of the involved risks in order to mitigate vulnerabilities and threats when decisions about dynamic collaboration are made. Moreover, a set of new metrics is defined to allow a novel form of risk quantification in these environments. Other contribu-

[3] http://www.espanavirtual.org/

[4] http://consequence.it.uc3m.es

tions of the paper include the definition of a generic hierarchical risk aggregation system, and a descriptive use-case where the risk computation framework is applied to enhance cloud-based service provisioning.

2. Title: *Enhancing Privacy and Dynamic Federation in IdM for Consumer Cloud Computing.*
   Authors: R. Sánchez, F. Almenares, P. Arias-Cabarcos, D. Díaz-Sánchez, A. Marín.
   Journal: IEEE Transactions on Consumer Electronics ISSN: 0098-3063. Printed version in Vol. 58, Iss. 1, 95 - 103, February 2012.(Impact Factor 2011: 0.941) [Sanchez et al., 2012].

   In this paper an architecture for dynamic federation with privacy improvements is presented. The document, extended from our conference paper in [Sánchez Guerrero et al., 2012], describes the first definition of the reputation assertion used in our trust model for conveying reputation data between providers.

3. Title: *FedTV: personal networks federation for IdM in mobile DTV.*
   Authors: F. Almenares, P. Arias-Cabarcos, D. Díaz-Sánchez, A. Marín, R. Sánchez.
   Journal: IEEE Transactions on Consumer Electronics. ISSN: 0098-3063. Printed version in Vol. 57, Iss. 2, 499 - 506, May 2011. (Impact Factor 2011: 0.941) [Almenarez et al., 2011].

   This work, extended from our conference paper in [Almenares et al., 2011], proposes an enhanced mobile client to support the establishment of federations to allow cooperation in mobile DTV scenarios (content sharing, service delegation, etc.). We extend the Enhanced Client Profile defined in SAML v2 [Cantor et al., 2005b], incorporating a trust management layer inside user's consumer electronic devices' software. Thus, the components of our architecture for dynamic federation are incorporated in a real world scenario that shows its benefits and applicability.

Furthermore, The following publications complement the core ideas in the above mentioned papers by the definition of application scenarios:

4. Title: *FamTV: An architecture for Presence-Aware Personalized Television.*
   Authors: P. Arias-Cabarcos, R. Sánchez, F. Almenares, D. Díaz-Sánchez, A. Marín.
   Journal: IEEE Transactions on Consumer Electronics. ISSN: 0098-3063. Printed version in Vol.57, no.1, pp.6-13, February 2011. (Impact Factor 2011: 0.941) [Arias

et al., 2011a].

This work, extended from our conference paper in [Arias et al., 2011b], presents a way to combine content adaptation paradigms together with presence detection in order to allow a seamless and personalized entertainment experience when watching TV. It includes a security layer where the trust-based federation will be used to dynamically cope with the huge ecosystem of services and applications that can be accesed from the TV. This work received the ***Chester W. Sall Award*** for the 2nd place best paper in the IEEE Transactions on Consumer Electronics 2011.

5. Title: *SuSSo: Seamless and Ubiquitous Single Sign-on for Cloud Service Continuity across devices.*
   Authors: P. Arias-Cabarcos, F. Almenares, R. Sánchez, A. Marín, D. Díaz-Sánchez.
   Journal: IEEE Transactions on Consumer Electronics. ISSN: 0098-3063. Printed version in Vol. 58, Iss. 4, 1425 - 1433, November 2012. (Impact Factor 2011: 0.941) [Arias et al., 2012c].

   This work, extended from our conference paper in [Arias et al., 2012a], presents an architecture for moving SAML sessions across devices guaranteeing cloud service continuity. It complements the proposal in [Almenarez et al., 2011] by tackling mobile scenarios. This work builds partly on the dynamic federation architecture modules presented in this thesis.

Main Contributions to International Conferences:

1. Title: *Towards dynamic trust establishment for identity federation.*
   Authors: F. Almenares, P. Arias-Cabarcos, A. Marín, D. Díaz-Sánchez.
   Conference: Euro American Conference on Telematics and Information Systems (EATIS 2009). Prague, Czech Republic, June 03 - 05, 2009 [Almenárez et al., 2009].

   This first paper analyzes the state-of-the-art and compares identity federation protocols. Based on the analysis, it identifies the need for new trust models that improve flexibility in federation scenarios, and outlines an initial conceptual description of the basic architectural requirements.

2. Title: *Enabling SAML for Dynamic Identity Federation Management.*

Authors: P. Arias-Cabarcos, F. Almenares, A. Marín, D. Díaz-Sánchez.
Conference: Wireless and Mobile Networking Conference. Gdansk, Poland, 2009 [Arias et al., 2009].

This paper completes the former by performing a deeper review of the existing identity federation frameworks, analyzing the underlying trust mechanisms and its suitability to be applied in open environments. Furthermore, we propose an extension for the SAML standard in order to facilitate the creation of federation relationships in a secure dynamic way between prior unknown parties based on the introduction of reputation. The realization of the approach, including a discussion of software components and a proof-of-concept implementation, is also described.

3. Title: *Risk Assessment for Better Identity Management in Pervasive Environments.*
Authors: P. Arias-Cabarcos.
Conference: IEEE PerCom Phd Forum 2011. Seattle, Washington, USA March 23, 2011 [Arias, 2011].

   The paper builds on the premise that risk evaluation must be considered as a key enabler to foster collaboration between parties in a dynamic but yet secure manner. The main idea outlined in the document is to enrich the decision making process in federated environments by introducing risk assessment and integrate it with trust evaluation, a solution not yet proposed in the published literature to that date. We first introduce the modeling of identity federation protocols as two-phased procedures (i.e., Pre-Federation and Post-Federation) and sketch a preliminary taxonomy to be used in determining risk metrics. This short paper was presented in a poster session during the PhD forum organized in the context of the 9th IEEE International conference on Pervasive Computing and Communications. The forum was structured as a combined one-day workshop prior to the conference and a poster session during the main conference sessions to encourage interaction between PhD students and researchers from academia, industry, and government. As a result of this participation, our contribution to the forum was awarded with the ***"Best PhD Forum Contribution Award"***.

4. Title: *Family Personalization Service.*
Authors: D. Díaz-Sánchez, R. Sánchez, P. Arias-Cabarcos, I. Bernavé, F. Almenares.
Conference: IEEE International Conference on Consumer Electronics - Berlin

(ICCE-Berlin 2011).

This work, builds on the ideas in [Arias et al., 2011b], and presents a personalization system that allows to automatically configure devices surrounding users. The system addresses privacy-based filtering and group preference modeling.

5. Title: *Introducing Infocards in NGN to enable user-centric identity management.*
Authors: D. Proserpio, F. Sanvido, P. Arias, R. Sánchez, D. Díaz-Sánchez, A. Marín, F. Almenares.
Conference: IEEE Global Communications Conference - Miami, Florida, USA (GLOBECOM 2010).

   This paper proposes a solution that leverages the benefits of the infoCard identity technology and introduces this user centric paradigm into the emerging NGN architectures.

1.4. Organization of the Thesis

In order to accomplish the goals outlined in the above sections, the organization of this dissertation is as follows:

Chapter 2 presents the state-of-the-art on technologies and latest research related to the thesis. It objectively reviews the different existing frameworks for identity management; provides an overview of trust/reputation and risk models; and summarizes related work being carried out by individual researchers, international research projects and organizations involved in standardization.

After this background, Chapter 3 goes into a deeper analysis of the research challenges in identity management and presents a detailed description of the research problem we aim to solve. The problem statement is clearly articulated based on this analysis and also the objectives pursued in this thesis are refined. The chapter ends with a value proposition, i.e., highlighting the potential impact of our research.

Chapter 4 proposes a generic infrastructure to solve the limitations of current identity federations. Based on this high-level infrastructure description, Chapters 5 and 6 go deeper into the main building blocks of the architecture, namely the the risk module and the trust and reputation module.

More specifically, in Chapter 5, we design a taxonomy to capture the different aspects of a relationship in identity federation that may contribute to risk. Based on the taxonomy and aiming at developing a computational model, we propose a set of metrics as a basis to quantify risk. We also describe how to aggregate the metrics into a meaningful risk figure, coming to the final formal definition of the model.

Next, Chapter 6 formalizes a trust model that captures the features of current FIM systems and introduces new trust dimensions to add flexibility and richness. We propose mechanisms to convey and use reputation data in the model.

After that, Chapter 7 explains how the trust and risk model proposals are integrated and used to make dynamic federation decisions; and Chapter 8 is dedicated to the validation of the ideas presented in this thesis.

Finally Chapter 9 summarizes the results and discussions presented in this thesis. Furthermore, since the need for further work and exploration is necessary in any useful research, we also describe the future lines that can be followed from the ideas presented here.

Apart from the the aforementioned chapters, we have included two appendices. Appendix A contains a glossary with all the acronyms used in the document; and Appendix B is a catalogue that summarizes the set of metrics for risk quantification proposed in the thesis.

Chapter 2

State of the art

"All men by nature desire knowledge."

Aristotle, On Man in the Universe

Contents

2.1. Federated Identity Management

Federated Identity Management, FIM, or Identity Federation [1] refer to the technologies, standards and use-cases which serve to enable the portability of identity information across otherwise autonomous security domains [Maler and Reed, 2008]. The ultimate goal of identity federation is to enable users of one domain to securely access data or systems of another domain seamlessly, and without the need for completely redundant user administration. The main actors in a FIM scenario are: 1) the Identity Provider (IdP), which vouches for the identity of a user and issues authentication, authorization and/or attribute tokens about her; 2) the Service Provider (SP)[2], which provides services to the end user and relies on the identity tokens generated by the IdP; and 3) the User, that interacts (usually via a user agent, e.g., web browser) with both SPs and IdPs. In this section we provide a general picture of the current identity landscape, explaining the existing federation protocols and specifications.

2.1.1. Security Assertion Markup Language (SAML)

Security Assertion Markup Language (SAML) is an standard developed by the Security Services Technical Committee (SSTC) of the standards organization OASIS (the Organization for the Advancement of Structured Information Standards).The latest version, SAMLv2.0 [Cantor et al., 2005b], became an OASIS Standard in March 2005 and it is built upon a number of existing standards, such as eXtensible Markup Language (XML) [Bray et al., 2008], XML Signature [Eastlake et al., 2012], XML Encryption [Eastlake et al., 2002b], Hypertext Transfer Protocol (HTTP) [Fielding et al., 1999], Simple Object Access Protocol (SOAP) [Box et al., 2000], etc. SAML defines a framework to allow the exchange of security information between online business partners. More specifically, it allows the exchange of authentication, attribute and/or authorization related data about a principal (usually an end user) between an Identity Provider and a Service Provider. Accordingly, Figure 2.1 shows the main concepts and components defined in the specifications.

Basically, as represented in the image in Figure 2.1, SAML specifies four different elements, which are detailed below:

[1] We will use these terms indistinctly throughout this thesis

[2] The term Relying Parties is also frequently used to refer to SPs

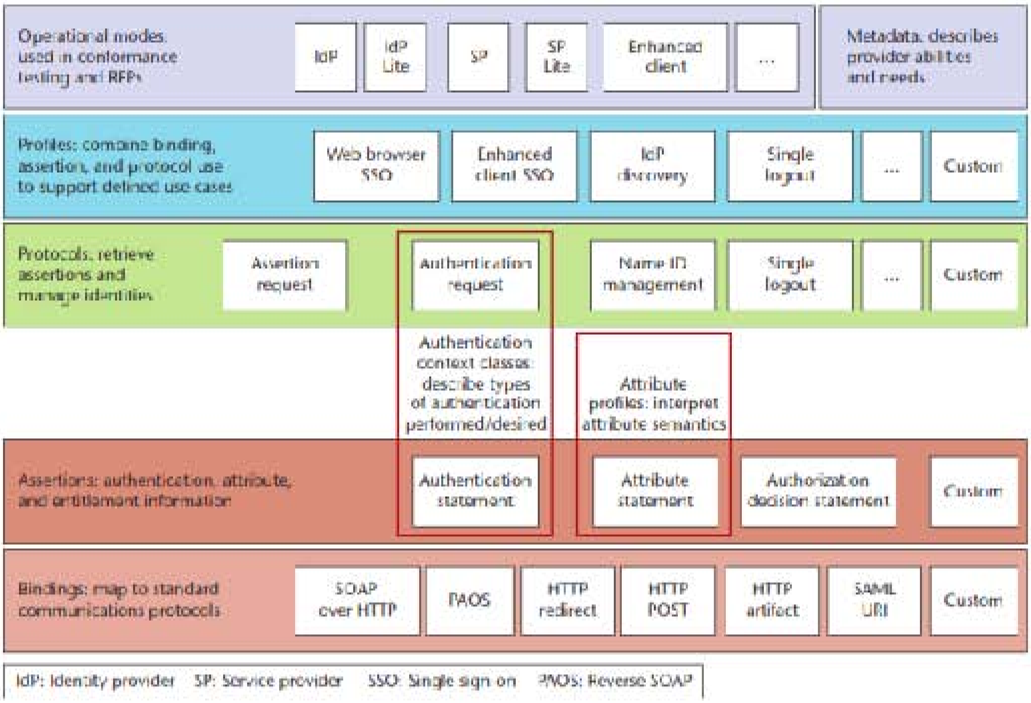

Operational modes used in conformance testing and RFPs
IdP
IdP Lite
SP
SP Lite
Enhanced client
...
Metadata: describes provider abilities and needs
Profiles: combine binding, assertion, and protocol use to support defined use cases
Web browser SSO
Enhanced client SSO
IdP discovery
Single logout
...
Custom
Protocols: retrieve assertions and manage identities
Assertion request
Authentication request
Name ID management
Single logout
...
Custom
Authentication context classes: describe types of authentication performed/desired
Attribute profiles: interpret attribute semantics
Assertions: authentication, attribute, and entitlement information
Authentication statement
Attribute statement
Authorization decision statement
Custom
Bindings: map to standard communications protocols
SOAP over HTTP
PAOS
HTTP redirect
HTTP POST
HTTP artifact
SAML URI
Custom
IdP: Identity provider SP: Service provider SSO: Single sign on PAOS: Reverse SOAP

Figure 2.1: The Security Assertion Markup Language framework (©[Maler and Reed, 2008]
).

- **Assertions** [Cantor et al., 2005b], which are statements related to authentication, attribute, or authorization about a subject, issued by an Identitiy Provider (IdP). The valid structure and contents of an assertion are defined by the SAML assertion XML schema. Regarding the kind of statements in the assertions:

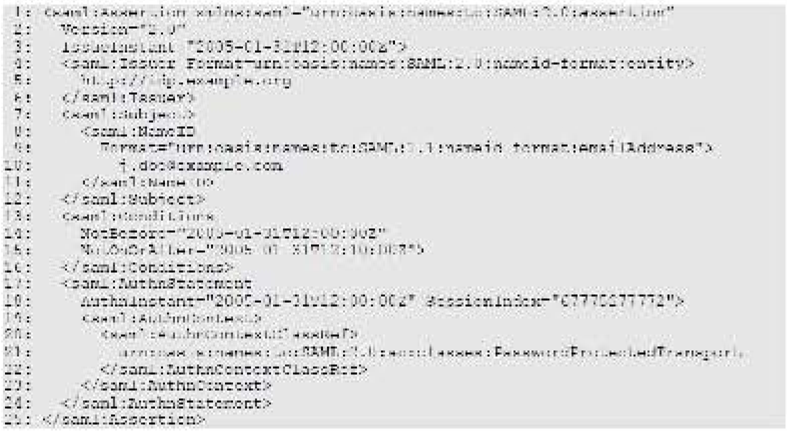
  - Authentication statements are created by the party that successfully authenticated a user. At a minimum, they describe the particular means used to authenticate the user and the specific time at which the authentication took place.

  - Attribute statements contain specific identifying attributes about the subject (e.g., that user "John Doe" has "Gold" card status).

  - Authorization decision statements define something that the subject is entitled to do (e.g., whether "John Doe" is permitted to buy a specified item).

  In order to better illustrate the format of a SAML Assertion, Figure 2.2 shows an

XML fragment containing an example assertion with a single authentication statement.



```
 1: <saml:Assertion xmlns:saml="urn:oasis:names:tc:SAML:2.0:assertion"
 2:   Version="2.0"
 3:   IssueInstant "2005-01-31T12:00:00Z">
 4:   <saml:Issuer Format=urn:oasis:names:SAML:2.0:nameid-format:entity>
 5:     http://idp.example.org
 6:   </saml:Issuer>
 7:   <saml:Subject>
 8:     <saml:NameID
 9:       Format="urn:oasis:names:tc:SAML:1.1:nameid-format:emailAddress">
10:         j.doe@example.com
11:     </saml:NameID>
12:   </saml:Subject>
13:   <saml:Conditions
14:     NotBefore="2005-01-31T12:00:00Z"
15:     NotOnOrAfter="2005-01-31T12:10:00Z">
16:   </saml:Conditions>
17:   <saml:AuthnStatement
18:     AuthnInstant="2005-01-31T12:00:00Z" SessionIndex="67775277772">
19:     <saml:AuthnContext>
20:       <saml:AuthnContextClassRef>
21:         urn:oasis:names:tc:SAML:2.0:ac:classes:PasswordProtectedTransport
22:       </saml:AuthnContextClassRef>
23:     </saml:AuthnContext>
24:   </saml:AuthnStatement>
25: </saml:Assertion>
```

Figure 2.2: Example fragment of a SAML assertion (©[Hughes and Maler, 2005]). SAML Assertions consist of XML packets containing information such as a target user's identifier, authentication status and attributes

The assertion in Figure 2.2 begins with the declaration of the SAML assertion namespace in line 1. Next, lines 2 through 6 provide information about the nature of the assertion: which version of SAML is being used, when the assertion was created, and who issued it. Lines 7 through 12 provide information about the subject of the assertion, to which all of the contained statements apply. The subject has a name identifier (line 10) whose value is "j.doe@example.com", provided in the format described on line 9 (email address). It can also be noted that the assertion as a whole has a validity period indicated by lines 14 and 15. Finally, the authentication statement appearing on lines 17 through 24 shows that this subject was originally authenticated using a password-protected transport mechanism (e.g. entering a username and password submitted over browser session protected with Secure Sockets Layer (SSL) [Freier et al., 2011]) at the time and date shown.

- **Protocols** [Cantor et al., 2005b], which define how and which assertions are requested. The set of SAML protocols and their descriptions are summarized in Table 2.1.

- **Bindings** [Cantor et al., 2005a], which define the lower-level communication or

Protocol	Description
Authentication Request Protocol	Defines a means by which a principal (or an agent acting on behalf of the principal) can request assertions containing authentication statements and, optionally, attribute statements
Single Logout Protocol	Defines a mechanism to allow near-simultaneous logout of active sessions associated with a principal. The logout can be directly initiated by the user, or initiated by an IdP or SP because of a session timeout, administrator command, etc.
Assertion Query and Request Protocol	Defines a set of queries by which SAML assertions may be obtained
Artifact Resolution Protocol	Provides a mechanism by which SAML protocol messages may be passed by reference using a small, fixed-length value called an artifact. The artifact receiver uses the Artifact Resolution Protocol to ask the message creator to dereference the artifact and return the actual protocol message
Name Identifier Management Protocol	Provides mechanisms to change the value or format of the name identifier used to refer to a principal. It also allows to terminate an association of a name identifier between an IdP and a SP
Name Identifier Mapping Protocol	Provides a mechanism to programmatically map one SAML name identifier into another, subject to appropriate policy controls

Table 2.1: SAML Protocols

messaging protocols (such as HTTP or SOAP) that the SAML protocols can be transported over. The set of SAML bindings and their descriptions are summarized in Table 2.2.

Binding	Description
HTTP Redirect Binding	Defines how SAML protocol messages can be transported using HTTP redirect messages (302 status code responses)
HTTP Post Binding	Defines how SAML protocol messages can be transported within the base64-encoded content of an HTML form control
HTTP Artifact Binding	Defines how an artifact is transported from a message sender to a message receiver using HTTP
SAML SOAP Binding	Defines how SAML protocol messages are transported within SOAP 1.1 messages, with details about using SOAP over HTTP
Reverse SOAP (PAOS) Binding	Defines a multi-stage SOAP/HTTP message exchange that permits an HTTP client to be a SOAP responder. Used in the Enhanced Client and Proxy Profile to enable clients and proxies capable of assisting in IdP discovery
SAML URI Binding	Defines a means for retrieving an existing SAML assertion by resolving a URI (Uniform Resource Identifier)

Table 2.2: SAML Bindings

- **Profiles** [Hughes et al., 2005], which are combinations of SAML protocols and bindings, together with the structure of assertions to cover specific use-cases. Profiles typically define constraints on the contents of SAML assertions, protocols, and bindings in order to solve the business use case in an interoperable fashion. The set of SAML bindings and their descriptions are summarized in Table 2.3.

Furthermore, there are two other SAML concepts defined in the specifications that are

Profiles	Description
Web Browser SSO Profile	Defines how SAML entities use the Authentication Request Protocol and SAML Response messages and assertions to achieve single sign-on with standard web browsers. It defines how the messages are used in combination with the HTTP Redirect, HTTP POST, and HTTP Artifact bindings
Enhanced Client and Proxy (ECP) Profile	Defines a specialized SSO profile where specialized clients or gateway proxies can use the Reverse-SOAP (PAOS) and SOAP bindings
Identity Provider Discovery Profile	Defines one possible mechanism for service providers to learn about the Identity Providers that a user has previously visited
Single Logout Profile	Defines how the SAML Single Logout Protocol can be used with SOAP, HTTP Redirect, HTTP POST, and HTTP Artifact bindings
Assertion Query/Request Profile	Defines how SAML entities can use the SAML Query and Request Protocol to obtain SAML assertions over a synchronous binding, such as SOAP
Artifact Resolution Profile	Defines how SAML entities can use the Artifact Resolution Protocol over a synchronous binding, such as SOAP, to obtain the protocol message referred to by an artifact
Name Identifier Management Profile	Defines how the Name Identifier Management Protocol may be used with SOAP, HTTP Redirect, HTTP POST, and HTTP Artifact bindings
Name Identifier Mapping Profile	Defines how the Name Identifier Mapping Protocol uses a synchronous binding such as SOAP

Table 2.3: SAML Profiles

useful for building and deploying a SAML environment:

- The **Metadata** [Cantor et al., 2005c], which define a way to express and share configuration information between SAML parties. For instance, these data can include an entity's support for given SAML bindings, identifier information, and Public Key Infrastructure (PKI) [Adams, C. and Farrell, S., 1999] information. SAML Metadata is defined by its own XML schema.

- The **Authentication Context** [Kemp et al., 2005], which can be used in SAML assertions to provide detailed information regarding the type and strength of authentication that a user employed when he authenticated at an identity provider. An SP can also include an authentication context in a request to an IdP to request that the user is authenticated using a specific set of authentication requirements, such as a multi-factor authentication. There is a general XML schema that defines the mechanisms for creating authentication context declarations and a set of SAML-defined "Authentication Context Classes", each with their own XML schema, that describe commonly used methods of authentication.

The combination of the aforementioned building-block components allow a number of use-cases to be supported. Arguably, the most important use case for which SAML is applied is multi-domain web Single Sign-on (SSO), shown in Figure 2.3 which allows a

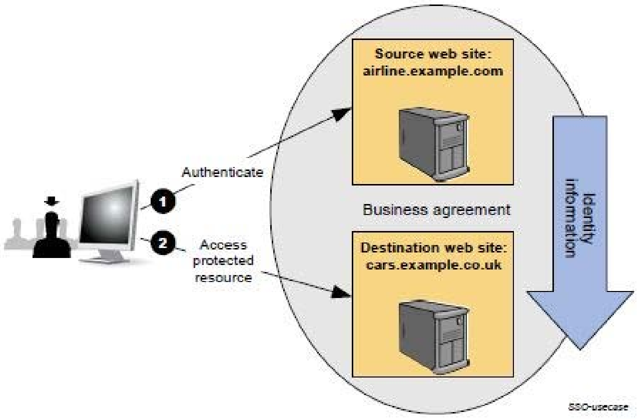


Figure 2.3: The Single Sign-On use case (©[Hughes and Maler, 2005]).

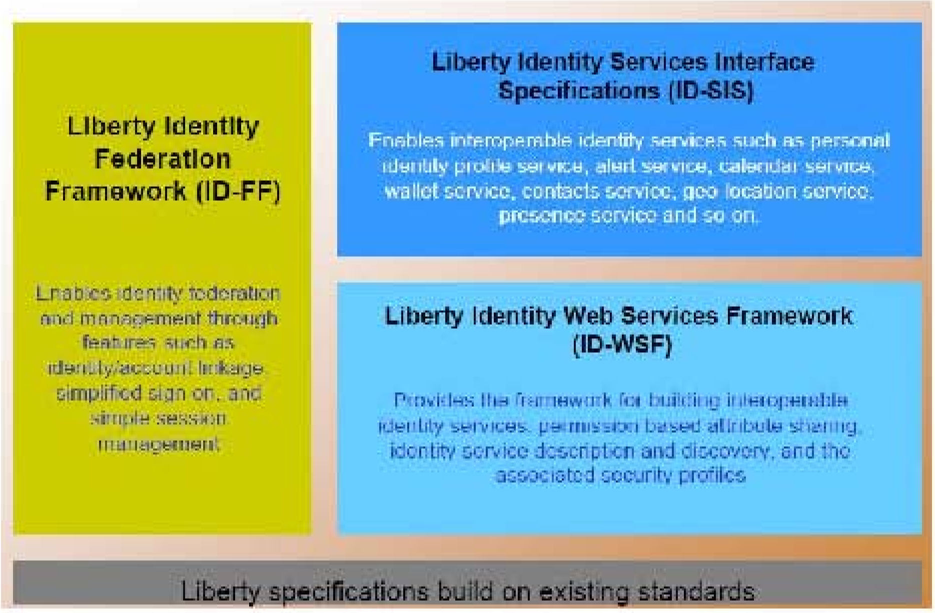
user to authenticate at a single site and gain access to other sites without the need for re-authentication, i.e., reusing the same identifier, act of authentication, and login session across multiple sites.

Apart from the aforementioned web SSO use case, SAML covers a huge range of use-cases, namely: Federation via Out-of-Band Account Linking, Federation via Identity Attributes, Federation via Transient Pseudonym or Persistent Identifiers and Federation Termination.

2.1.2. Liberty Alliance

Liberty Alliance (LA) [Project, 2012] was formed in September 2001 by a group of organizations with the aim to establish open standards to easily conduct online transactions while protecting the privacy and security of identity information. Based on this philosophy, the Liberty project designed an architecture and a set of protocols that provide support for federated identity management. The development of the core Liberty specifications was organized in three phases, as shown in Figure 2.4:

- **Phase 1- Liberty Identity Federation Framework (ID-FF):** In July 2002, the Liberty Alliance released its first set of public specifications, Liberty Identity Federation (ID-FF) 1.0. At this time, several member companies also announced upcoming availability of Liberty-enabled products, marking very rapid release and deployment of open specifications. The Liberty Alliance released two more versions of the Identity Federation specification, and then in June 2003 contributed its federation



Liberty Identity Federation Framework (ID-FF)

Enables identity federation and management through features such as identity/account linkage, simplified sign on, and simple session management

Liberty Identity Services Interface Specifications (ID-SIS)

Enables interoperable identity services such as personal identity profile service, alert service, calendar service, wallet service, contacts service, geo location service, presence service and so on.

Liberty Identity Web Services Framework (ID-WSF)

Provides the framework for building interoperable identity services, permission based attribute sharing, identity service description and discovery, and the associated security profiles

Liberty specifications build on existing standards

Figure 2.4: Liberty Alliance core frameworks (©[Liberty Alliance, 2013])

specification, to OASIS, forming the foundation for SAML 2.0 [3].

The ID-FF framework [Cantor and (eds.), 2003] defines a set of protocols, bindings, and profiles that provides a solution for identity federation, cross-domain authentication, and session management. This framework can be used to create a new identity management system or to develop one in conjunction with legacy systems. ID-FF is designed to work with heterogeneous platforms, various networking devices (including personal computers, mobile phones, and personal digital assistants), and emerging technologies. ID-FF is built upon the concept of Circle of Trust (CoT), which is defined by Liberty as a federation of SPs and IdPs that have business relationships based on Liberty architecture and operational agreements and with whom users can transact business in a secure and apparently seamless environment.

Summarizing, the Liberty ID-FF Protocols and Schema Specifications define transmission formats for the functions explained in Table 2.4.

- **Phase 2- Liberty Identity Web Services Framework (ID-WSF):** Liberty Al-

[3] Though some differences between SAML V2.0 and ID-FF V1.2 exist (http://saml.xml.org/differences-between-saml-v2-0-and-liberty-id-ff-1-2), we will use the terms SAML and Liberty throughout this thesis to generally refer to SAML-based FIM frameworks

ID-FF Protocol	Function Description
Single Sign-On and Federation Protocol	Defines the rules for request and response messages with which a principal is able to authenticate to one or more service providers and federate (or link) configured identities.
Name Registration Protocol	Defines the request and response messages a service provider would use to create its own opaque handle to identify a principal when communicating with the identity provider. This registration would occur after federation has been accomplished. After the service provider registers this new handle, subsequent communications with the identity provider would use this identifier rather than the identifier originally defined by the identity provider.
Federation Termination Notification Protocol	Defines a one-way message that one provider would use to notify another provider when a principal has terminated identity federation.
Single Logout Protocol	Defines the request and response messages that providers would exchange when notifying each other of logout events. This exchange would terminate all sessions when a logout occurs at either the service provider or the identity provider.
Name Identifier Mapping Protocol	Defines the request and response messages that one service provider can use to communicate with a second service provider to obtain the name identifier assigned to a principal federated in the name space of the second service provider.

Table 2.4: Functions covered by ID-FF protocols

liance also focused on identity web services standards, publicly releasing the Liberty Identity Web Services Framework in April 2004. Liberty ID-WSF [Beatty et al., 2004] provides the framework for building interoperable identity services, permission based attribute sharing, identity service description and discovery, and the associated security profiles.

The aim is at providing specifications for identity-based web services to work in tandem with the previous Liberty ID-FF. Thus, the Liberty ID-WSF can be used to develop web services that retrieve, update, or perform an action on identity data in a federated network environment using a SOAP-based invocation. ID-WSF introduces three new subjects apart from the defined in the ID-FF specifications, namely:

- A Web Service Consumer (WSC), which invokes the functions provided by a web service by making a request to the web service's provider.
- A Web Service Provider (WSP), which implements a web service based on a request from a WSC.
- A Discovery Service (DS), which permits: a) registration of services associated with an identity (i.e., each WSP registers the identity service that hosts to a DS); and b) discovery of services associated with an identity (i.e., the WSC queries DS in order to retrieve WSP data)

- **Phase 3- Liberty Identity Services Interface Specifications (ID-SIS):** Since

2003 Liberty also worked on this set of specifications that enable interoperable identity services. Thus, the Liberty Identity Service Interface Specifications [Kellomaki and Wason, 2003] comprise a set of identity services built on top of the ID-WSF framework. These services, which are supposed to have strong demand by the industry, are:

- Personal Profile, which describes a web identity service that provides a principal's basic profile information, such as their contact details, or name.
- Employee Profile, which describes a web identity service that provides a employee's basic profile information, such as their contact details, or name.
- Contact book, which describes a web identity service that allows a principal to manage contacts for private and business relations, and for the principal himself.
- Geolocation, which specifies a web identity service offering geolocation information associated with a principal.
- Presence, which specifies a web identity service offering presence information associated with a principal.
- Directory Access Protocol, which describes a web service offering directory information as an instance of a data-oriented identity web service, based on the Liberty ID-WSF data services template.
- Content SMS and MMS, which describes a web service that layers the ID-WSF 1.1 framework on the Multimedia Messaging Service interface type MM7 to add identity-based invocation and addressing.

Apart from the core frameworks defined above, during the last years of the project Liberty Alliance also released two more frameworks dealing with governance and identity assurance issues, namely:

- **Identity Governance Framework (IGF):** In February 2007, the Liberty Alliance started to work on the Identity Governance Framework [Madsen, 2009], releasing the first version publicly in July 2007. The Identity Governance Framework defines a set of standards to help enterprises easily determine and control how identity related information is used, stored, and propagated in appropriate and secure ways.

- **Identity Assurance Framework (IAF):** The Liberty Alliance started to work on the Identity Assurance Framework [Cutler, 2007] in 2008. This framework details four identity assurance levels designed to ease and speed the process of linking trusted identity-enabled enterprise, social networking and Web 2.0 applications together based on standardized business rules and security risks associated with each level of identity assurance. The assurance levels are based on the four levels of assurance outlined by the U.S. National Institute of Standards and Technology (NIST) Special Publication 800-63 version 1.0.1 [Nadalin et al., 2006], and range in confidence level from "low" to "very high". The level of assurance provided is measured by the strength and rigor of the identity proofing process, the credential's strength, and the management processes the service provider applies to it. These four assurance levels have been adopted by the U.K. government, the Government of Canada and the U.S. Federal Government for categorizing electronic identity trust levels for providing electronic government services.

Since 2009, the work of the Liberty Alliance is transitioning to the Kantara Initiative [4] , which means that all the Liberty Alliance material has been contributed to this new organization. Kantara is a non-profit professional association dedicated to advancing technical and legal innovation related to digital identity management. It is not a standards body but it submits recommendations to standards bodies such as OASIS, Internet Engineering Task Force (IETF), International Organization for Standardization (ISO), International Telecommunication Union Standardization Section (ITU-T) and other standard-developing organizations. As defined in the Kantara web site, the mission of the initiative is to foster identity community harmonization, interoperability, innovation, and broad adoption through the development of open identity specifications, operational frameworks, education programs, deployment and usage best practices for privacy-respecting, secure access to online services

2.1.3. WS-Federation

The Web Service Federation Language or WS-Federation [Goodner and (eds.), 2009] is an OASIS standard that forms part of the larger Web Services Security framework (WS-*). More specifically, WS-Federation describes how to use WS-Trust [Nadalin et al., 2009],

[4] http://kantarainitiative.org

WS-Security [Nadalin et al., 2004] and WS-Policy [World Wide Web Consortium (W3C), 2007] all together in order to provide federation between security domains. This enables high value scenarios where authorized access to resources managed in one realm can be provided to security principals whose identities and attributes are managed in other realms. WS-Federation includes mechanisms for brokering of identity, attribute discovery and retrieval, authentication and authorization claims between federation partners, and protecting the privacy of these claims across organizational boundaries.

WS-Federation relies on the Security Token Service (STS) model defined by WS-Trust, and a protocol (involving Request Security Token (RST) and Response messages) for handling such tokens, which contain information described by WS-SecurityPolicy. The STS is used to broker an establishment of a trust relationship between resource providers / relying parties and other service providers. Different federation services can be developed as variations of the base STS. Furthermore, processing in WS-Federation is kept independent of the security token format and the type of token being transmitted. WS-Federation defines also a metadata model and a document format describing how services can be discovered and combined, as well as their access policies. The types of services in WS-Federation are:

- **Authorization services:** can be viewed as decision brokering services. Interoperability of services requires a common model for interacting with authorization services.
- **Authentication type services:** a set of URIs is defined for specifying the parameter that sets the type of authentication in Request Security Token and Response messages.
- **Attribute services:** WS-Federation defines a model for accessing attribute services which may be needed to establish a federation context, e.g., information for advanced functionality or personalized user experience.
- **Pseudonym services:** allow principals to have different aliases in different realms or for different resources. They provide different kinds of identity mappings, e.g., with pseudonyms established per login or per service. In combination with the attribute services, they allow information to be provided about a requestor identified by a pseudonym, if the requestor has authorized this.

- **Privacy services:** extensions to WS-Trust syntax are defined to express both privacy requirements of a requester and mechanisms used by a STS for issuing a token. This may include, e.g., identification of sensitive claims in a token that must be protected by encryption.

2.1.4. OpenID

OpenID is defined as an open, decentralized, and free framework for user-centric digital identity management. It is based on well-known existing Internet technologies (HTTP, SSL, Diffie-Hellman [Rescorla, 1999]), and it is clearly oriented to be used in web scenarios. The efforts to develop OpenID started in 2005 and the framework is currently defined in a set of open specifications [OpenID, 2007] [Hardt et al., 2007] [Recordon et al., 2008] [Recordon and Fitzpatrick, 2006] [Hoyt et al., 2006].

The protocol operation for user authentication is easy: basically an identity is represented by means of an URI and the authentication process involves verifying that the user owns this URI. According to it, when a user wants to log into an OpenID enabled web site (or Relying Party), the browser (or User Agent) is redirected to the OpenID Identity Provider (OP), who attempts to authenticate the user and informs the web site of its success or failure. A major feature of OpenID is user-centricity, which means that users can decide which Identity Provider they trust the most to authenticate them. In fact, users can also become their own IdP without the need of registration or authorization from a third party. Thus, the OpenID protocol does not rely on a central authority to authenticate a user's identity.

The OpenID Authentication protocol flow is depicted in Figure 2.5. It consists of the following steps:

- In steps 1, 2 and 3 the end user tries to access a web site and, after being presented with the OpenID prompt page, he initiates authentication by presenting a user-supplied identifier to the Relying Party via their User-Agent.

- In step 4, after normalizing the user-supplied identifier, the Relying Party performs discovery on it and establishes the OP endpoint URL that the end user uses for authentication.

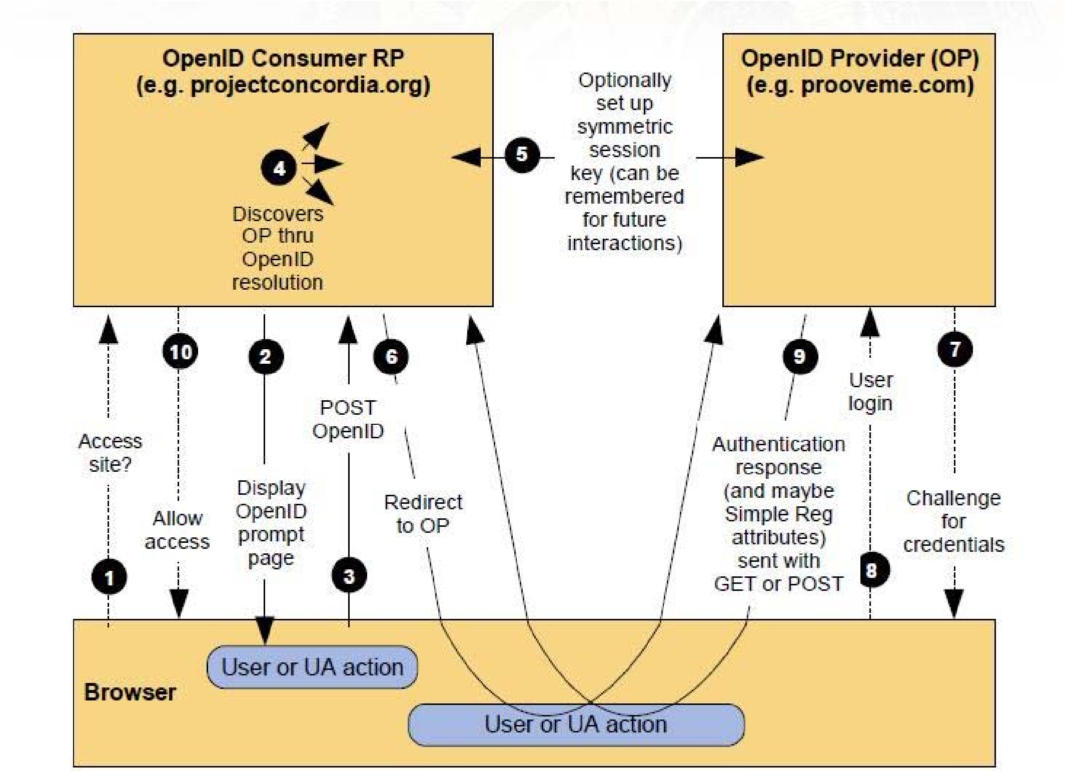

OpenID Consumer RP
(e.g. projectconcordia.org)
4
Discovers OP thru OpenID resolution
5
Optionally set up symmetric session key (can be remembered for future interactions)
OpenID Provider (OP)
(e.g. prooveme.com)
10
2
6
9
User login
7
Access site?
POST OpenID
Display OpenID prompt page
Allow access
Redirect to OP
Authentication response (and maybe Simple Reg attributes) sent with GET or POST
Challenge for credentials
1
3
8
Browser
User or UA action
User or UA action

Figure 2.5: OpenID Authentication protocol flow

- Optionally in step 5 the Relying Party and the OP establish an association – a shared secret established using Diffie-Hellman Key Exchange. The OP uses an association to sign subsequent messages and the Relying Party to verify those messages; this removes the need for subsequent direct requests to verify the signature after each authentication request/response.

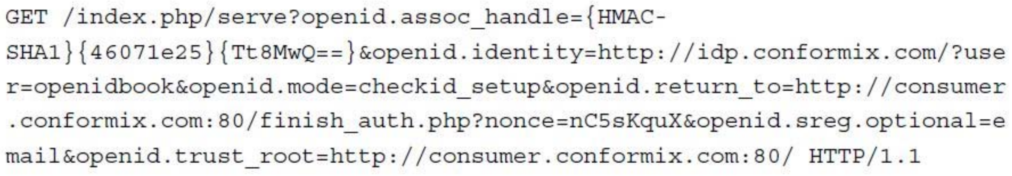
- In step 6 the Relying Party redirects the end user's User-Agent to the OP with an OpenID Authentication request.

- In steps 7 and 8 the OP establishes whether the end user is authorized to perform OpenID Authentication and wishes to do so.

- In step 9 the OP redirects the end user's User-Agent back to the Relying Party with either an assertion that authentication is approved or a message that authentication failed.

- In step 10 the Relying Party verifies the information received from the OP and decides whether to allow access to the service or not.

The OpenID Authentication protocol messages are mappings of plain-text keys to plain-text values exchanged using the HTTP protocol. In order to generally illustrate the message format, Figure 2.6 shows an example of OpenID authentication request message (step 6).



```
GET /index.php/serve?openid.assoc_handle={HMAC-
SHA1}{46071e25}{Tt8MwQ==}&openid.identity=http://idp.conformix.com/?use
r=openidbook&openid.mode=checkid_setup&openid.return_to=http://consumer
.conformix.com:80/finish_auth.php?nonce=nC5sKquX&openid.sreg.optional=e
mail&openid.trust_root=http://consumer.conformix.com:80/ HTTP/1.1
```

Figure 2.6: Example of OpenID Authentication Request message

As it can be seen in Figure 2.6, an OpenID message contains a series of keys (the openid message parameters) preceded with the "openid." prefix. For each key a value is provided and the whole message is codified to be sent over an HTTP GET request.

Summarizing, OpenID is mainly an authentication protocol and federation is achieved with extensions, such as [Hardt et al., 2007] that allows attribute exchange.

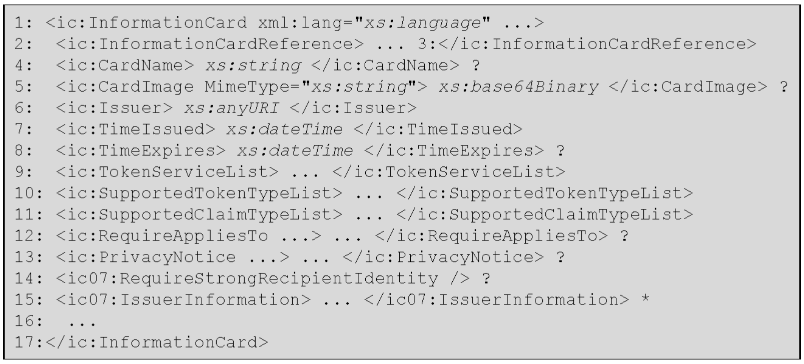
2.1.5. Information Cards

The Information Card (aka infoCard) technology allows to represent personal digital identities that people can use online. Conceptually, Information Cards are the digital version of the physical cards we carry in our purse or wallet today. In line with this metaphor, Information Cards are handled by users with a new kind of "digital wallet" called a selector.

The Information Card technology is advanced by a non-profit organization - the Information Card Foundation (IFC)[5] - composed by companies and individuals working together to evolve the internet identity ecosystem. The foundation is currently organized in a series of active Working Groups that deal with issues such as standardization, implementation guidelines and best practices, interoperability with deployed identity technologies, etc. Furthermore, they published several white papers and specifications on Information Card technology and practice, being [Nanda and Jones, 2008] and [Jones and (eds.), 2009] the core documents that define the identity formats and protocol flows.

[5] http://informationcard.net

As defined in the aforementioned specifications, the identity information is represented as a signed XML document, also called security token.



```
<ic:InformationCard xml:lang="xs:language" ...>
  <ic:InformationCardReference> ... 3:</ic:InformationCardReference>
  <ic:CardName> xs:string </ic:CardName> ?
  <ic:CardImage MimeType="xs:string"> xs:base64Binary </ic:CardImage> ?
  <ic:Issuer> xs:anyURI </ic:Issuer>
  <ic:TimeIssued> xs:dateTime </ic:TimeIssued>
  <ic:TimeExpires> xs:dateTime </ic:TimeExpires> ?
  <ic:TokenServiceList> ... </ic:TokenServiceList>
  <ic:SupportedTokenTypeList> ... </ic:SupportedTokenTypeList>
  <ic:SupportedClaimTypeList> ... </ic:SupportedClaimTypeList>
  <ic:RequireAppliesTo ...> ... </ic:RequireAppliesTo> ?
  <ic:PrivacyNotice ...> ... </ic:PrivacyNotice> ?
  <ic07:RequireStrongRecipientIdentity /> ?
  <ic07:IssuerInformation> ... </ic07:IssuerInformation> *
  ...
</ic:InformationCard>
```

Figure 2.7: Information Card format

Figure 2.7 shows the XML schema for Information Cards. As it can be noted, it contains information such as expiration time, issuer data, type of supported claims, etc. The key data in this document are the claims or user attributes. Depending on who is vouching for the claims, two types of Information Cards are specified: 1) Personal or Self-Issued cards, which represent a small, fixed attribute set whose values are determined solely by the user (e.g., phone number, e-mail address, web address); and 2) Managed Information Cards, which contain claims issued by Identity Providers. The latter can be auditing, non-auditing, or auditing-optional to accommodate the needs of different business models. Based on the low level XML data, each Information Card has a visual representation in the form of a card-shaped picture and a card name associated with it. This graphic metaphor enables users to organize their digital identities and to easily select one they want to use for any given interaction. The participants in digital identity interactions are IdPs, SPs and users, with the particularity that users interact through an Identity Selector.

The Identity Selector is an active client that allows users to store, manage, and use their digital identities. Examples of identity selectors are Microsoft's Windows CardSpace [Bertocci et al., 2007], and several kinds of Identity Selectors from the Eclipse Higgins Project [Higgins Project, 2009]. Among the key functionalities of a selector, the most remarkable ones are: providing a consistent user experience for authentication based on a graphic interface, allowing the creation and managing of personal Information Cards, and facilitating the import and export of Information Cards in standard file formats.

The diagram in Figure 2.8 depicts the basic protocol flow when using an Information

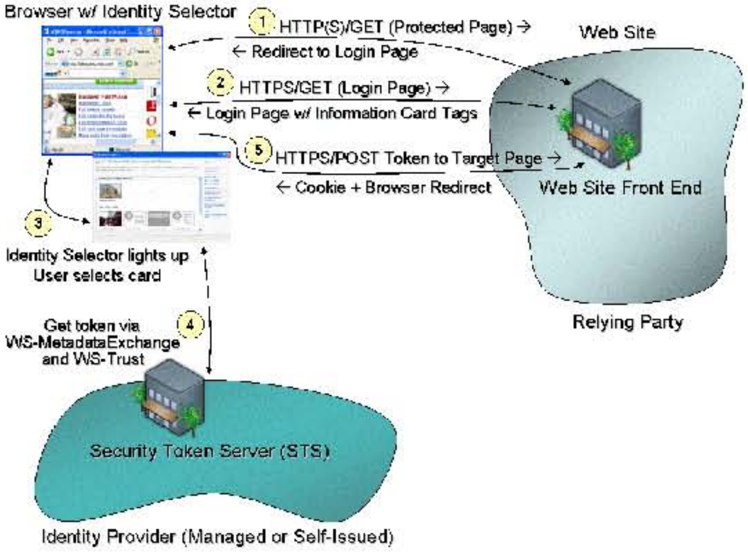

Browser w/ Identity Selector
1 HTTP(S)/GET (Protected Page) →
← Redirect to Login Page
Web Site
2 HTTPS/GET (Login Page) →
← Login Page w/ Information Card Tags
5 HTTPS/POST Token to Target Page →
← Cookie + Browser Redirect
Web Site Front End
3
Identity Selector lights up
User selects card
Relying Party
Get token via
WS-MetadataExchange
and WS-Trust
4
Security Token Server (STS)
Identity Provider (Managed or Self-Issued)

Figure 2.8: Basic protocol flow when using an Information Card at a web site (©[Jones and (eds.), 2009])

Card at a web site. Steps 1, 2, and 5 are essentially the same as a typical forms-based login today: (1) The user navigates to a protected page that requires authentication. (2) The site redirects the browser to a login page, which presents a Web form. (5) The browser posts the Web form that includes the login credentials supplied by the user back to the login page. The site then validates the contents of the form including the user credentials, typically writes a client-side browser cookie to the client for the protected page domain, and redirects the browser back to the protected page. The key difference between this scenario and today's site login scenarios is that the login page returned to the browser in step (2) contains an HTML tag that allows the user to choose to use an Information Card to authenticate to the site. When the user selects this tag, the browser invokes an Identity Selector, which implements the Information Card user experience and protocols, and triggers steps (3) through (5). In Step (3), the browser Information Card support code invokes the Identity Selector, passing it parameter values supplied by the Information Card HTML tag supplied by the site in Step (2). The user then uses the Identity Selector to choose an Information Card, which represents a digital identity that can be used to authenticate at that site. Step (4) retrieves a Security Token that represents the digital identity selected by the user from the Security Token Service at the Identity Provider for that identity. In Step (5), the browser posts the token obtained back to the Web site using a HTTPS/POST. The web site validates the token, completing the user's

Information Card-based authentication to the Web site. Following authentication, the web site typically then writes a client-side browser cookie and redirects the browser back to the protected page.

Summarizing, the key feature of Information Card based identity systems is the focus on a user centric experience. In addition, Information Cards support several data formats and authentication methods such as SAML, and OpenID. On the other hand, beyond being used to log into web sites, Information Cards can also facilitate other kinds of interactions based on attribute exchange. One possible use of claims is online age verification, with Identity Providers providing proof-of-age cards, and Relying Parties accepting them for purposes such as online wine sales; other attributes could be verified as well. Another is online payment, where merchants could accept online payment cards from payment issuers, containing only the minimal information needed to facilitate payment. Furthermore, role statements carried by claims can be used for access control decisions by Relying Parties.

2.1.6. O-Auth

OAuth is an open-source specification for authorization, which requires implicit federation. It defines a framework for allowing a third-party application (the “Consumer” or “Client”) to access protected resources from another application (the “Service Provider”, or “Resource Owner”) at the request of a “User” of the Client application. OAuth allows the user to enter his user credentials (e.g., username and password) only to the provider, which then grants the Client permission to view the protected resources on behalf of the user.

Though it is mainly a delegation protocol, the federation model is supported. The OAuth specification simply assumes that there is some form of authentication mechanism in place that is acceptable to the SP. It could be local authentication (e.g., as seen on Facebook, etc), or federation from SAML, OpenID, etc.

The OAuth 1.0 Protocol was published as RFC 5849 [Hammer-Lahav, 2010], an informational Request for Comments, in April 2010. Currently, OAuth 2.0 is a work in progress at the IETF. This evolution of the initial version focuses on client developer simplicity while providing specific authorization flows for web applications, desktop applications, mobile phones, and living room devices.

The OAuth Authorization scheme uses the standard HTTP Authorization and WWW-Authenticate headers to pass OAuth Protocol Parameters. According to the specifications, Table 2.5 shows the main OAuth concepts and definitions.

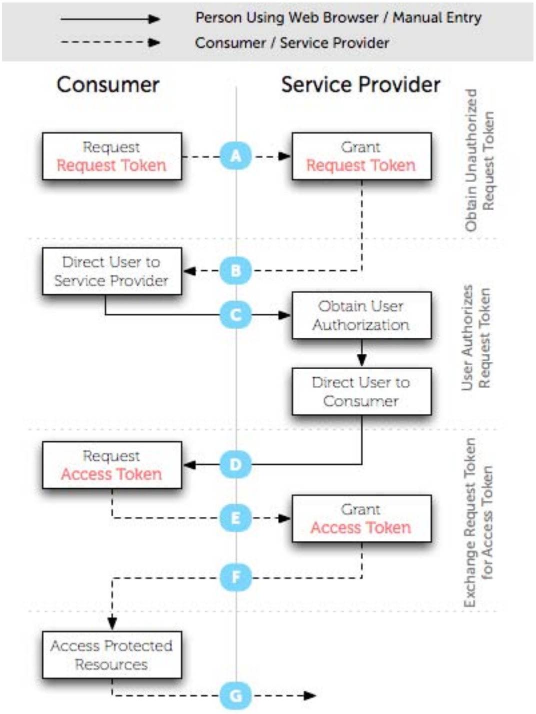
OAuth concept	Definition
Service Provider	A web application that allows access via OAuth
User	An individual who has an account with the Service Provider
Consumer	A website or application that uses OAuth to access the Service Provider on behalf of the user
Protected Resource(s)	Data controlled by the Service Provider, which the Consumer can access through authentication
Consumer Key	A value used by the Consumer to identify itself to the Service Provider
Request Token	A value used by the Consumer to obtain authorization from the user, and exchanged for an Access Token
Access Token	A value used by the Consumer to gain access to the Protected Resources on behalf of the user, instead of using the user's Service Provider credentials
Token Secret	A secret used by the Consumer to establish ownership of a given Token.

Table 2.5: Relevant OAuth concepts

An example OAuth use case is allowing printing service `printer.example.com` (the Consumer), to access private photos stored on `photos.example.net` (the Service Provider) without requiring users to provide their `photos.example.net` credentials to `printer.example.com` . The underlying OAuth protocol flow that would take place in a scenario like the mentioned example is depicted in Figure 2.9.

The steps in the protocol are explained below:

- In step A the Consumer asks for a Request Token by sending an HTTP request to the Service Provider's Request Token URL. The request must be signed and contain, among other parameters, the Consumer Key that identifies the Consumer.
- In step B The Service Provider verifies the signature and Consumer Key. If successful, it generates a Request Token and Token Secret and returns them to the Consumer in the HTTP response body.
- In step C, the Consumer must obtain approval from the user by directing the user to the Service Provider.
- In step D the Service Provider verifies the user's identity and asks for consent. After the user authenticates with the Service Provider and grants permission for Consumer access, the Consumer must be notified that the Request Token has been authorized



Person Using Web Browser / Manual Entry
Consumer / Service Provider
Consumer
Service Provider
Request Request Token
A
Grant Request Token
Obtain Unauthorized Request Token
Direct User to Service Provider
B
C
Obtain User Authorization
Direct User to Consumer
User Authorizes Request Token
Request Access Token
D
E
Grant Access Token
F
Exchange Request Token for Access Token
Access Protected Resources
G

Figure 2.9: OAuth v1.0 Authentication flow (©[Atwood et al., 2007])

and ready to be exchanged for an Access Token. Thus, the Service Providers directs the user back to the Consumer.

- In step E, the Consumer makes an HTTP request to the Service Provider's Access Token URL in order to obtain an Access Token. The request must be signed and contains, among other parameters, the Consumer Key and the Request Token previously obtained.

- In step F, the Service Provider veryfies the request and, if successful, generates an Access Token and a Token Secret that are returned in the HTTP response body.

- In step G, the Consumer is able to access the protected resources on behalf of the user by generating signed requests that contain the Access Token and Token Secret granted in the previous step.

In summary, OAuth aims to unify the experience and implementation of delegated web service authentication into a single, community-driven protocol. The specification builds on existing protocols and best practices that have been independently implemented by various websites. An open standard, supported by large and small providers alike, promotes a consistent and trusted experience for both application developers and the users of those applications.

2.2. Trust and Reputation Models

2.2.1. Basic Concepts and Definitions

Trust and reputation are present in our daily lives and constitute an important basis for security since they are indispensable to allow cooperation between strangers. These concepts have been widely investigated in a range of disciplines and academic domains, such as psychology, economy or sociology. Similarly, since the advent of the digital era, trust and reputation emerged as vital concepts also in the field of computer science. The significance of incorporating trust and reputation systems to this field lies on the fact that they are enabling technologies which aid in decision making, support secure online transactions, and whose inclusion is expected to guarantee the long-term growth and success of the Internet [The Internet Society, 2008].

For these reasons, there is a rapidly growing literature around trust and reputation systems, being [McKnight and Chervany, 1996], [Jøsang et al., 2007], [Sabater and Sierra, 2005] and [Gómez Mármol and Martínez Pérez, 2010] representative samples of research on the topics. More specifically, [McKnight and Chervany, 1996] is a classical paper that elaborates on the meanings of trust; [Jøsang et al., 2007] and [Sabater and Sierra, 2005] are surveys of trust and reputation models; and [Gómez Mármol and Martínez Pérez, 2010] is a more recent work that extracts the common points in trust and reputation models and provide a series of guidelines towards standardization.

Thus, based on the literature, it can be stated that computational trust tries to apply the human notion of trust into the digital world with the aim to increase the reliability and performance of electronic communities. However, due to the subjective nature and the applicability of the terms in different contexts and from different perspectives, there is a lack of consensus in the definition of trust. It can be observed that trust is an abstract and complex notion related to concepts such as e.g., confidence, reliance, dependence, or faith. For these reasons, trust is quite challenging to formalize and many definitions have been given. To name a few relevant ones:

- Gambetta states that "*trust (or, symmetrically, distrust) is a particular level of the subjective probability with which an agent will perform a particular action, both before [we] can monitor such action (or independently of his capacity of ever to be able to monitor it) and in a context in which it affects [our] own action*" [Gambetta, 2000].

- Jøsang sees trust as "*a belief that one entity has about another entity. Firstly, there must be a reason behind the belief, and secondly, the belief expresses an expectation of how an entity will behave or perform*" [Jøsang, 1996].

- As defined by Marsh trust is "*a useful judgment in the light of experience of the behavior of others*" [Marsh, 1994].

- According to the ITU-T, "*generally an entity can be said to trust a second entity when the first entity makes the assumption that the second entity will behave exactly as the first entity expects*" [ITU, 2000].

Despite the lack of a unique consistent definition, a number of commonly identified properties of trust can be derived from the above statements and the vast number of definitions found in the literature. In this sense, trust is usually specified in terms of a relationship between a *trustor*, the subject that trusts a target entity, which is known as the *trustee* i.e., the entity that is trusted. Trust forms the basis for allowing a trustee to use or manipulate resources owned by a trustor or may influence a trustor's decision to use a service provided by a trustee. Thus, trust can form an important factor in decision-making. Furthermore, a number works distinguish between situational versus general trust, being the first associated to particular situations or contexts and the latter a general measure of the global trustworthiness of an entity independently of the context. Trust ultimately is a personal and subjective phenomenon that is based on various factors or evidences, and that some of those carry more weight than others. For example, personal experience typically carries more weight than second hand trust referrals, but in the absence of personal experience, trust often has to be based on referrals from others.

On the other hand, reputation has also been widely studied in the literature. In this case, the existing definitions show that the main property inherent to the concept of reputation is the sense of "collective thinking". Therefore, as described in [Jøsang et al., 2007], reputation can be considered as a "*collective measure of trustworthiness (in the sense of reliability) based on the referrals or ratings from members in a community*".

In reputation systems, usually four distinct types of agents or roles are involved [Casare and Sichman, 2005]. Despite the terminology may vary, the semantics of these roles are:

- Evaluators: these are agents who can develop an evaluation or evaluative belief about other agents, including individuals, groups, organizations, etc. The information used by evaluators can be direct experiences with the targets or through third parties.
- Targets: these are agents that play the role of the evaluation object.
- Beneficiaries: these are individuals, groups, organizations, etc., who benefit from the evaluation.
- Propagators: these are third parties that can propagate the reputation information to other agents who need the information, usually beneficiaries.

Furthermore, according to Resnik et al. [Resnick and Zeckhauser, 2002] reputation systems need to have the following three properties to operate:

- Longevity of agents: agents are long lived, which means that it should be impossible for an agent to change his/her identity or pseudonym to erase the records about his/her past behaviors. Without longevity, agents can erase their bad reputation scores easily, so new reputation scores may not reflect their real reputation status.
- Protocol of ratings: reputation systems need to have a certain protocol by which ratings about current interactions are captured and distributed.
- Usability of reputation system: ratings about past interactions must be useful to guide certain decisions or actions. There is no reason for reputation systems to exist without any usability.

However, because the notions of trust and reputation themselves are vague, what constitutes a trust or a reputation system is difficult to describe concisely. Nevertheless, it is clear that there are important differences between the two concepts. Basically, trust systems produce a score that reflects the relying party's subjective view of an entity's trustworthiness, whereas reputation systems produce an entity's public reputation score as seen by the whole community. This means that an entity can trust others based on their good reputation, while it can also trust some other entity with a bad reputation because it has a certain knowledge based on past direct experience or referral relationships. But trust and reputation models also have certain key processes in common (as identified in [Gómez Mármol and Martínez Pérez, 2010]) such as scoring, ranking, rewarding, punishing or gathering behavioral information.

There can of course be trust systems that incorporate elements of reputation systems and vice versa, so that it is not always clear how a given systems should be classified. In fact, trust and reputation are used interchangeably in some of the existing literature, since the use of reputation information may foster trust.

Here, we provide a brief summary of these models based on the network architecture, as done in [Jøsang et al., 2007]: centralized or distributed systems. For each class, a general description of the model operation is provided together with the main limitations and advantages, and some examples of existing models in the category and their application scenarios.

2.2.2. Centralized Models

In a centralized reputation system, all the information about the performance of a given participant is collected as ratings from other members in the community who have had direct experience with that participant. In these systems, a central authority is in charge of collecting the ratings, computing a reputation score for every participant and making these scores publicly available. Hence, participants can access each other's scores and use them when deciding whether or not to transact with a particular entity. Fig. 2.10 shows a typical centralized reputation system, where A and B denote transaction partners with a history of transactions in the past, and who consider transacting with each other in the present.

The global reputations are updated after every transaction, since transacting entities provide ratings about each other's performance in the transaction.

Centralized reputation systems have important advantages. Firstly, the protocol for gathering/conveying reputation data is easier. Every entity just has to communicate with the central authority to submit votes or retrieve reputation about others in a client-server manner. Secondly, the computation of aggregated values is performed by the central server so there is no computation overhead in the clients.

On the other hand, there are also some drawbacks. Since the architecture is centralized,

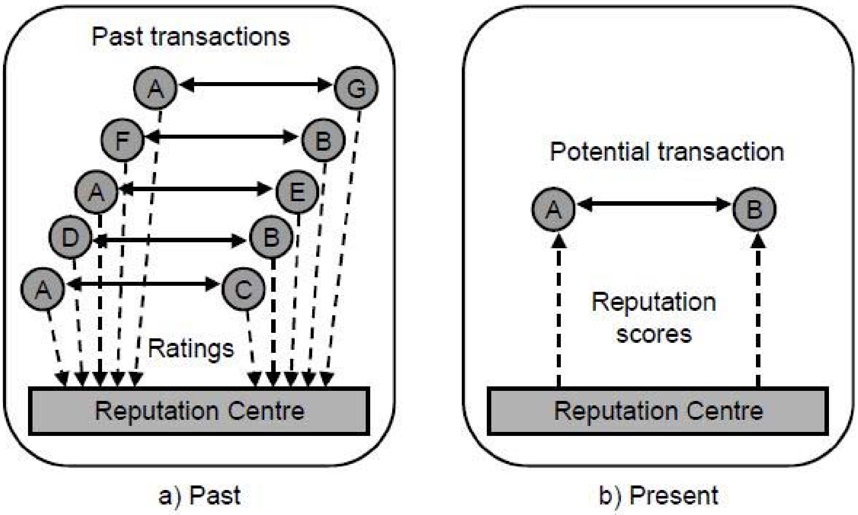


Figure 2.10: Centralised reputation system (©[Jøsang et al., 2007])

the reputation server becomes a single point of failure, as well as a bottleneck which can restrict the flow of transactions. In addition, the scalability of this systems is limited. On the one hand it may be unfeasible to have a single authority that all entities trust. And on the other hand, as the number of entities grow, providing performance and robustness requires a large invest of money for the central authority.

Due to their simplicity, most of the successfully deployed reputation systems have a centralized architecture. Among popular well-known systems, we can cite EBay [6], the Epinions system [7] for products and shop reviews, or Google's Web Page Ranking System (PageRank) [Page et al., 1999].

2.2.3. Distributed Models

In a distributed reputation system there is no central authority for submitting ratings or obtaining reputation scores of others. Instead, different distributed approaches can be followed. There can be distributed stores where ratings can be submitted, or each entity simply records the opinion about experiences with other parties, and provides this information on request. Thus, whenever an entity wants to transact with another unknown party, it has to first find the distributed stores, or try to obtain opinions from entities that

[6] http://www.ebay.com
[7] http://www.epinions.com

have had direct experience with that target party. Fig. 2.11 shows the operation model of a distributed reputation system.

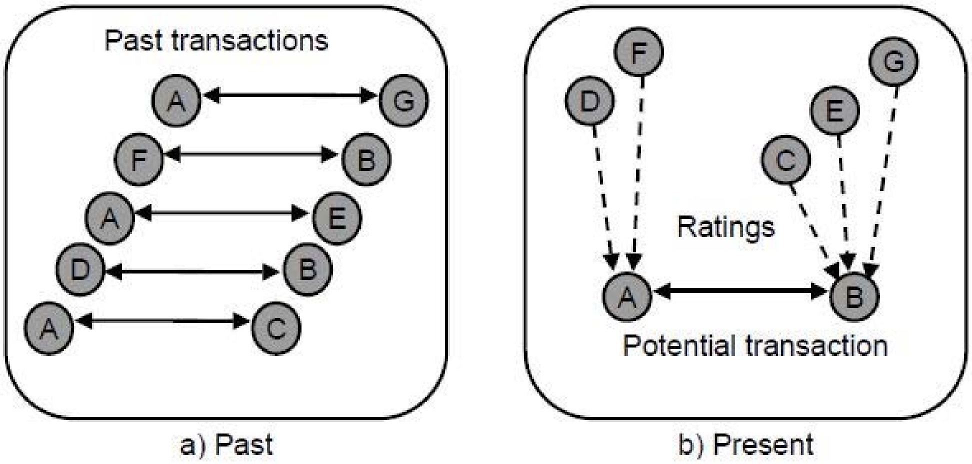


Figure 2.11: Distributed reputation system (©[Jøsang et al., 2007])

After obtaining the scores from the distributed stores or peers, it is the asking entity the one that has to compute the aggregated reputation value based on the received ratings. Obviously, if the entity has had direct experience with the target party, it can be also taken into account for the calculation, possibly carrying a higher weight than the received opinions. Since there is no central authority this kind of system do not have a single point of failure. However, the process of data dissemination and reputation calculation gets more complex. In a distributed environment, each participant is responsible for collecting and combining ratings from other participants. In this conditions it is often impossible or too costly to obtain ratings resulting from all interactions with a given party. Instead the reputation score is based on a subset of ratings.Peer-to-Peer (P2P) networks, where every node plays the role of both client and server, represent an environment well suited for distributed reputation management.

The application of trust and reputation systems to P2P networks is specially useful to identify unreliable or malicious participants in order to isolate and avoid transactions with them. Thus, many authors have proposed trust and reputation models for P2P networks, such as the Eigentrust algorithm [Kamvar et al., 2003], approaches [Cornelli et al., 2002] and [Damiani et al., 2002] to identify reputable servents and reliable resources in P2P applications for file exchange, or the PTM [Almenárez et al., 2004] model designed to manage trust relationships between peers in open and dynamic environments.

Summarizing, centralized reputation systems rely on a central entity to gather, compute and disseminate reputation information. Distributed reputation systems on the other hand rely on decentralized solutions where every peer stores information about the other peers with which they interacted, and dissemination is performed on demand between peers. To conclude, centralized reputation systems work well and are easier to deploy. However, there are environments where a distributed reputation system, i.e., without any centralized functions, is better suited than a centralized one.

2.3. Risk Assessment

2.3.1. Basic Concepts and Definitions

Risk is a very general concept and it has been interpreted in different ways depending on the specific application. Consequently, there are many definitions in the literature, sometimes inconsistent and ambiguous.

Here we name a few definitions that help in contextualizing this thesis:

- Risk is defined as ' '*the possibility of something bad happening*" in the Cambridge dictionary[8].
- As defined in the wikipedia[9], "*risk concerns the expected value of one or more results of one or more future events. Technically, the value of those results may be positive or negative.*"
- According to the ISO Guide 73:2002 [Guide, ISO, 2002], risk is "*the effect of uncertainty on objectives.*"
- In the norm OHSAS (Occupational Health and Safety Assessment Series) 18001:2007 [Palomino and Rivero, 2008] risk is defined as "*a combination of the likelihood of an occurrence of a hazardous event or exposure (s) and the severity of injury or ill health that can be caused by the event or exposure (s).*"

[8] dictionary.cambridge.org

[9] en.wikipedia.org

- According to NIST SP 800-30 [Stoneburner et al., 2002], risk is "*a function of the likelihood of a given threat-source's exercising a particular potential vulnerability, and the resulting impact of that adverse event on the organization.* "

The common theme in these definitions and in those found in the literature is a concern with potential future harm or loss. Thus, most definitions view risk as a combination of the probability of an undesired event and the magnitude of the impact if it occurs. Risk is strictly tied to uncertainty. Uncertainties include events (which may or not happen) and uncertainties caused by ambiguity or a lack of information.

More formally (and quantitatively), risk is proportional to both the results expected from an event and to the probability of this event. Mathematically, risk is generally defined as:

$$Risk = Probability\ of\ Event * Impact\ of\ Event \tag{2.1}$$

Since risk is crucial for secure decision making, it is useful to create models for risk assessment that allow to calculate risk based on the generic formula in (2.1). Risk assessment can be defined as the determination of a risk value for a specific context, which is a step of risk management. Risk management is the identification, assessment, and prioritization of risks followed by coordinated and economical application of resources to minimize, monitor, and control the probability and/or impact of unfortunate events or to maximize the realization of opportunities [Purdy, 2010].

The risk assessment step can be done following different approaches: quantitative analysis, semi-quantitative analysis and qualitative analysis. [ENISA, 2006]. In quantitative analysis numerical values are assigned to both impact and likelihood. These values are derived from a variety of sources. Impact can be determined by evaluating and processing the various results of an event or by extrapolation from experimental studies or past data. The quality of the entire analysis depends on the accuracy of the assigned values and the validity of the statistical models used. The advantage of this analysis is that it provides a precise numerical risk value which is useful for cost benefit analysis of recommended controls. However, depending on the numerical ranges used to express the measurement, the meaning of the numerical risk value may lead to ambiguities; a high risk value can be due to the high value of the asset or the high probability of loss or both factors. Thus, high risks due to high probability and low impact may be considered equal to high risks due to low probability and high impact, and the meaning is different in each case.

On the other hand, the qualitative analysis uses descriptive variables to represent the magnitude and likelihood of potential consequences. The scales used can be formed or adjusted to suit the circumstances, and different descriptions may be used for different risks. This kind of analysis is useful for example when non-tangible aspects of risk are to be considered (e.g., reputation, culture, image, etc.) or when there is a lack of adequate information and numerical data or resources necessary for a statistically acceptable quantitative approach. This kind of analysis provides a mean to identify and assess risks in a relatively shorter time. However, the cost benefit analysis of recommended controls becomes difficult in the absence of a precise numerical risk value.

Finally, in semi-quantitative analysis approaches the objective is to assign numeric values to the scales used in the qualitative assessment. The mapping of these values to the risk can be obtained using a mapping table based on the advises of the security experts. These values are usually indicative and not real, which is the prerequisite of the quantitative approach. Therefore, as the value allocated to each scale is not an accurate representation of the actual magnitude of impact or likelihood, the numbers used must only be combined using a formula that recognizes the limitations or assumptions made in the description of the scales used.

As risk carries so many different meanings there are also a number of formal methods used to assess or to “measure” risk. Here we summarize four well-known methodologies in order to provide a brief background that contextualizes the thesis.

NIST 800-30 Risk Management Guide for Information Technology Systems and CVSS

The NIST 800-30 Risk Management Guide for Information Technology Systems [Stonebburner et al., 2002] is a risk management methodology developed by the National Institute for Standards and Technology. The steps of risk analysis using NIST 800-30, summarized in Figure 2.12, are described below.

1. System Characterization. This step involves the collection of system-related information of different kinds (hardware, software, system interfaces, etc.) The goal is to establish the scope of the risk management efforts.

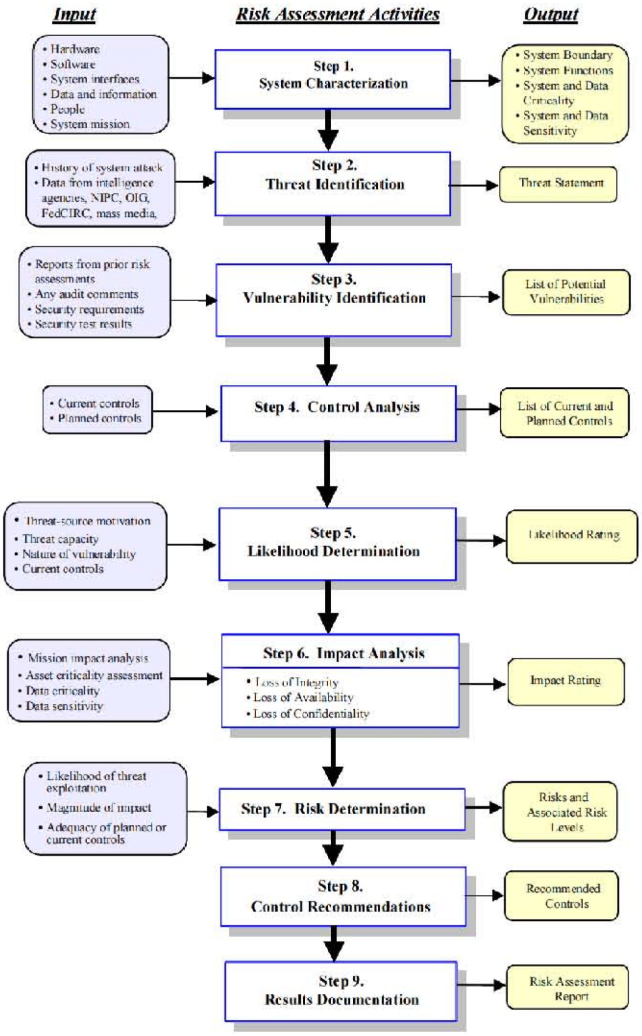


Figure 2.12: NIST 800-30 risk management methodology (©[Stoneburner et al., 2002])

2. Threat Identification. The goal of this step is to identify the potential threat sources and compile a threat statement listing potential threat sources that are applicable to the system being evaluated.

3. Vulnerability Identification. The goal of this step is to develop a list of system vulnerabilities (flaws or weaknesses) that could be exploited by the potential threat sources.

4. Control Analysis. The goal of this step is to analyze the controls that have been implemented, or are planned for implementation, by the organization to minimize or eliminate the likelihood (or probability) of a threat's exercising a system vulner-

ability.

5. Likelihood Determination. This step involves the mapping of vulnerabilities to their associated likelihood. The likelihood that a potential vulnerability could be exercised by a given threat-source can be described as High, Medium, or Low. This likelihood levels are defined in a descriptive qualitative manner in the scale Low, Medium and High.

6. Impact Analysis. The goal of this step is to determine the adverse impact resulting from a successful threat exercise of a vulnerability. The adverse impact of a security event can be described in terms of loss or degradation of any, or a combination of any, of the following three security goals: integrity, availability, and confidentiality.

   Some tangible impacts can be measured quantitatively in lost revenue, the cost of repairing the system, or the level of effort required to correct problems caused by a successful threat action. Other impacts (e.g., loss of public confidence, loss of credibility, damage to an organizationŠs interest) cannot be measured in specific units but can be qualified or described in terms of High, Medium, and Low impacts. Because of the generic nature of this discussion, the NIST 800-30 guide designates and describes only the qualitative categories High, Medium, and Low impact.

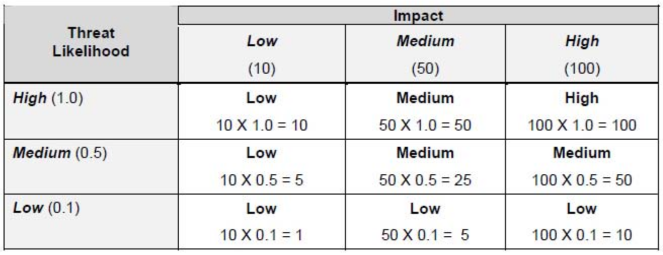
7. Risk Determination. The purpose of this step is to assess the level of risk. The determination of risk for a particular threat/vulnerability pair can be expressed as a function of: (1) the likelihood of a given threat-source's attempting to exercise a given vulnerability, (2) the magnitude of the impact should a threat-source successfully exercise the vulnerability, and (3) the adequacy of planned or existing security controls for reducing or eliminating risk. To measure risk, a risk-level matrix such as the sample matrix depicted in Figure 2.13 must be developed. Thus, the final determination of risk is derived by multiplying the ratings assigned for threat likelihood and threat impact. The sample matrix in Figure 2.13 shows how the overall risk levels of High, Medium, and Low are derived. The determination of these risk levels or ratings may be subjective. The rationale for this justification can be explained in terms of the probability assigned for each threat likelihood level and a value assigned for each impact level. In this example:

   - The probability assigned for each threat likelihood level is 1.0 for High, 0.5 for

Medium, 0.1 for Low.

- The value assigned for each impact level is 100 for High, 50 for Medium, and 10 for Low.

And the final risk scale assigns High Risk for values >50 to 100; Medium Risk for values >10 to 50; and Low Risk for values from 1 to 10.



Threat Likelihood	Impact		
	Low (10)	*Medium* (50)	*High* (100)
High (1.0)	Low 10 X 1.0 = 10	Medium 50 X 1.0 = 50	High 100 X 1.0 = 100
Medium (0.5)	Low 10 X 0.5 = 5	Medium 50 X 0.5 = 25	Medium 100 X 0.5 = 50
Low (0.1)	Low 10 X 0.1 = 1	Low 50 X 0.1 = 5	Low 100 X 0.1 = 10

Figure 2.13: Risk Level Matrix sample according to NIST 800-30 methodology (©[Stoneburner et al., 2002])

8. Control Recommendations. During this step of the process, controls that could mitigate or eliminate the identified risks, as appropriate to the organizationŠs operations, are provided. The goal of the recommended controls is to reduce the level of risk to the IT system and its data to an acceptable level.

9. Results Documentation. Once the risk assessment has been completed (threat-sources and vulnerabilities identified, risks assessed, and recommended controls provided), the results should be documented in an official report.

Magerit

Magerit was prepared and promoted by CSAE (Consejo Superior de Administración Electrónica), Spain in response to the perception that the government (and, in general, the whole society) increasingly depends on information technologies for achieving its service objectives [Ministerio de Administraciones Públicas de España, 1999]. The Magerit methodology fo risk analysis is comprised of the following five steps:

1. Determination of assets. This step involves identifying those assets that are relevant for the organization, their interrelationships and their value. The essential asset is the information handled by the system, i.e., the data, but other relevant assets can

be identified around these data, for example the services that can be provided to these data or the computer applications that allow these data to be handled.

2. Determination of threats. The goal of this step is to determine the threats that may affect each asset. Threats are "*things that happen.*" Of all the things that could happen, those that are of interest are those that could happen to our assets and cause damage. There are threats from natural disasters (earthquakes, floods, etc) and industrial accidents (pollution, electrical failures, etc). There are threats caused by persons, either through errors or intentional attacks.

3. Determination of safeguards. The goal of this step is to determine what safeguards are available and how effective they are against the risk. Safeguards or countermeasures are procedures or technological mechanisms that reduce the risk.

4. Determination of the impact. This step involves the estimation of the impact, defined as the measurement of the damage to an asset arising from the occurrence of a threat.

5. Determination of the risk. This step involves the estimation of the risk, defined as the measurement of the probable damage to the system. Risk is the weighted impact on the rate of occurrence (or the expectation of appearance) of the threat. Knowing the impact of the threats to the assets, the risk can be derived by taking into account the frequency of occurrence.

OWASP Risk Rating Methodology

The risk rating methodology proposed by OWASP (The Open Web Application Security Project) [OWASP, 2012] is based on the conception of risk as equal to *Likelihood x Impact*. They first decompose the likelihood and impact in factors and then show how to combine these factors to determine the overall severity for the risk.

The methodology is comprised of six steps:

1. *Risk identification.* This step consists of identifying a security risk that needs to be rated. It means gathering information about the threat agents involved, the attack they are using, the vulnerability involved, and the impact of a successful exploit.

2. *Break down of factors for estimating likelihood.* At the highest level, this is a rough measure of how likely a particular vulnerability is to be exploited by an attacker. The

factors that influence the estimate of the likelihood are divided into two categories: *threat agent factors* and *vulnerability factors*. Each category has a set of options and each of the options has a likelihood rating from 0 to 9 (see Table 2.6), which will be used to estimate the overall likelihood.

Factor Category	Option	Rating
Threat Agent Factors	Skill level	How technically skilled is this group of threat agents? No technical skills (1), some technical skills (3), advanced computer user (4), network and programming skills (6), security penetration skills (9)
	Motive	How motivated is this group of threat agents to find and exploit this vulnerability? Low or no reward (1), possible reward (4), high reward (9)
	Opportunity	What resources and opportunity are required for this group of threat agents to find and exploit this vulnerability? full access or expensive resources required (0), special access or resources required (4), some access or resources required (7), no access or resources required (9)
	Size	How large is this group of threat agents? Developers (2), system administrators (2), intranet users (4), partners (5), authenticated users (6), anonymous Internet users (9)
Vulnerability Factors	Ease of discovery	How easy is it for this group of threat agents to discover this vulnerability? Practically impossible (1), difficult (3), easy (7), automated tools available (9)
	Ease of exploit	How easy is it for this group of threat agents to actually exploit this vulnerability? Theoretical (1), difficult (3), easy (5), automated tools available (9)
	Awareness	How well known is this vulnerability to this group of threat agents? Unknown (1), hidden (4), obvious (6), public knowledge (9)
	Intrusion Detection	How likely is an exploit to be detected? Active detection in application (1), logged and reviewed (3), logged without review (8), not logged (9)

Table 2.6: Contributing factors for likelihood estimation according to OWASP Risk Rating Methodology [OWASP, 2012]

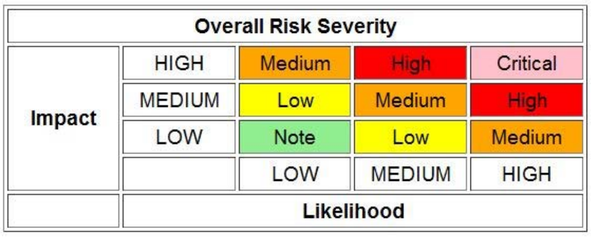
3. *Break down of factors for estimating impact.* There are two kinds of impact factors. The first is the *technical impact* on the application, the data it uses, and the functions it provides. The other is the *business impact* on the business and company operating the application. As in the case of likelihood estimation, each factor has a set of options, and each option has an impact rating from 0 to 9 (see Table 2.7) associated with it. This ratings will be used to estimate the overall impact.

4. *Determining severity of the Risk.* In this step the overall risk is calculated. For this purpose both likelihood and impact are computed as the average of the scores of each option. After calculating these two numerical values, thay are classified as

Factor Category	**Option**	**Rating**
Technical Factors	Loss of confidentiality	How much data could be disclosed and how sensitive is it? Minimal non-sensitive data disclosed (2), minimal critical data disclosed (6), extensive non-sensitive data disclosed (6), extensive critical data disclosed (7), all data disclosed (9)
	Loss of integrity	How much data could be corrupted and how damaged is it? Minimal slightly corrupt data (1), minimal seriously corrupt data (3), extensive slightly corrupt data (5), extensive seriously corrupt data (7), all data totally corrupt (9)
	Loss of availability	How much service could be lost and how vital is it? Minimal secondary services interrupted (1), minimal primary services interrupted (5), extensive secondary services interrupted (5), extensive primary services interrupted (7), all services completely lost (9)
	Loss of accountability	Are the threat agents' actions traceable to an individual? Fully traceable (1), possibly traceable (7), completely anonymous (9)
Bussiness Factors	Financial damage	How much financial damage will result from an exploit? Less than the cost to fix the vulnerability (1), minor effect on annual profit (3), significant effect on annual profit (7), bankruptcy (9)
	Reputation damage	Would an exploit result in reputation damage that would harm the business? Minimal damage (1), Loss of major accounts (4), loss of goodwill (5), brand damage (9)
	Non-compliance	How much exposure does non-compliance introduce? Minor violation (2), clear violation (5), high profile violation (7)
	Privacy violation	How much personally identifiable information could be disclosed? One individual (3), hundreds of people (5), thousands of people (7), millions of people (9)

Table 2.7: Contributing factors for impact estimation according to OWASP Risk Rating Methodology [OWASP, 2012]

LOW, MEDIUM or HIGH. Less than 3 is LOW, 3 to less than 6 is MEDIUM, and 6 to 9 is HIGH. Finally, likelihood and impact are combined according to Figure 2.14 in order to get a final severity rating for the risk.



Overall Risk Severity				
Impact	HIGH	Medium	High	Critical
	MEDIUM	Low	Medium	High
	LOW	Note	Low	Medium
		LOW	MEDIUM	HIGH
	Likelihood			

Figure 2.14: Overall risk calculation according to OWASP Risk Rating Methodology (©[OWASP, 2012])

5. *Deciding what to fix.* After following the previous steps, the outcome is a prioritized list of risks to fix. OWASP recommends to fix the most severe risks first in order to improve the overall risk profile. It is also important to take into account the cost associated to implementing controls to fix a risk, since some risks mighth not be

worth fixing.

6. *Customizing the risk rating model.* Finally, OWASP recognizes that there is not a risk rating methodology that is universaly applicable. Thus, costumization is allowed and the following mechanisms are considered as convenient ways to tailor the model:

   - Adding factors: different factors that better fit the applicacion can be chosen. For example, a military application might add impact factors related to loss of human life or classified information.

   - Customizing options: the provided options are just a general sample, but new options can be added. Furthermore, the scores associated with the options can be also changed.

   - Weighting factors: the model above assumes that all the factors are equally important. However, the factors can be weighted to emphasize those that are more significant for a spcific context.

Despite the various methodologies in conducting security risk analysis and them being tailored to particular contexts, there are commonalities in their steps. Basically, all the methodologies share a common framework similar to the following procedure: (1) assets, vulnerabilities and threats identification; (2) risk assessment; (3) selection of controls; and (4) re-evaluation.

We have chosen NIST 800-30, Magerit and OWASP as representative examples of risk management methods, but more risk methodologies have been developed. The survey in [ENISA, 2006] provides a more detailed compilation of risk methodologies.

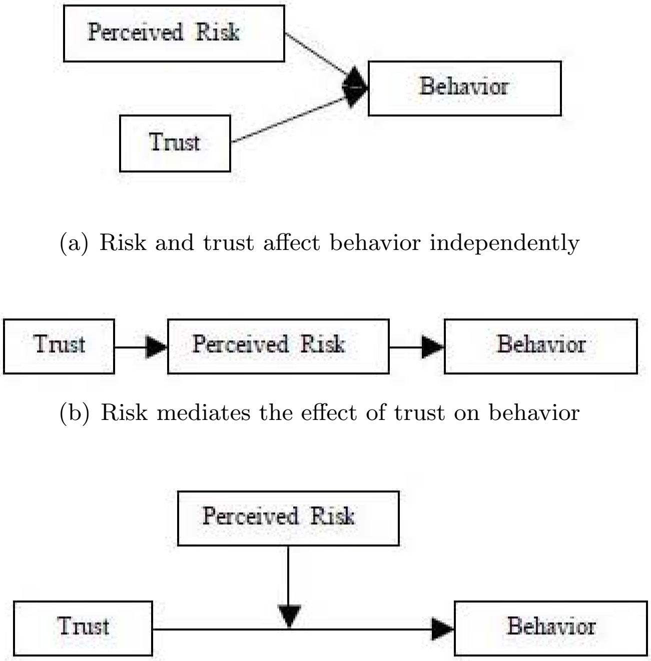
As stated before, risk is tied to uncertainty. But there is also another concept closely related to risk: trust. A decision to trust is usually associated with an explicit or implicit assessment of risk. For example, if risk is low, it is easier to trust; but if risk is high, trust is generally less willingly assumed. Nevertheless, the relationship between trust and risk is much more complex and therefore hard to formalize. In the next section, a literature review of works considering risk evaluation within trust models is provided.

2.3.2. Risk Considerations in Trust Models

The relationship between risk and trust concepts has also been widely studied in the literature. In "*Why Trust is not proportional to Risk*" [Solhaug et al., 2007], Solhaug et al. remark that it is crucial to understand the relationship between trust and risk in order to allow secure trust-based cooperation and they also provide an analysis in this regard.

They criticize Josang and Presti's idea in [Jøsang and Presti, 2004] that "*the more trustworthy a potential partner, the less risk involved in doing transactions with him.*" The authors also disagree with the trust view presented by Grandison and Sloman [Grandison and Sloman, 2000], which suggests that "*the level of trust has an approximate inverse relationship to the degree of risk*". In turn, their vision is that trust is generally neither proportional nor inverse proportional to risk. Rather, they state that as higher trustworthiness means lower probability of an incident, trust is inverse to the probability of a risk and proportional to the value the trustee is willing to stake, i.e., proportional to the consequence of a risk. The risk can hence not be determined from the trust value alone. As mentioned in [Solhaug et al., 2007], there is a number of other influential contributions to the area of trust and trust management that are unclear about the precise relation between trust and risk. Motivated by the lack of consistency in the literature, Gefen et al. [Gefen et al., 2003] highlight also the need for clarification in the relationship between trust and risk concepts. They point out that the IT literature on the topic embraces three primary models for the relationship between trust and risk. These three models, depicted in Figure 2.3.2, are: (1) the consideration of trust and risk as independent factors; (2) the consideration of a mediating relationship between both factors; and (3) the consideration of a moderating relationship.

The first case encompasses those studies which hypothesize that trust and risk are not related in a specific cause-effect relation. Instead, the vision about the trust/risk relationship according to this model is that both simultaneously affect behavior in an independent way. On the other hand, the mediating relationship model suggested in a number of research works states that trust affects perceived risk, which affects behavior. Thus, the existence of trust reduces the perception of risk, which in turn increases the willingness to transact. Finally, the last model conceptualizes the relationship between trust and risk as a moderating relationship, i.e., it is believed that the effect of trust on behavior is different when the level of risk is low versus when the level of risk is high.



(a) Risk and trust affect behavior independently

(b) Risk mediates the effect of trust on behavior

(c) Risk moderates the effect of trust on behavior

Figure 2.15: Models of trust and risk relationship (©[Gefen et al., 2003])

2.4. Related Work

As will be shown in Chapter 3, none of the above identity management solutions define a suitable trust and risk model to allow dynamic and secure federation establishment. Or, equivalently, no trust/risk model has been yet defined that is appropriate for the specific federation scenario. This section reviews both the related work carried out by researchers (individually or in the framework of a research project), as well as the standardization initiatives that are related or may contribute in any aspect to realize the vision of dynamic federation.

2.4.1. Individual Research and International Projects

The Internet2 group, Ping Identity and Stockholm University are working in "*Distributed Dynamic SAML*" [Internet2, 2008] to deal with challenges regarding deployment, scalability and interoperability of SAML-based federation deployments. They aim to achieve: 1) distribution, in the sense of changing the operations of typical multi-party SAML federations to be less dependent on central administration; and 2) dynamism, which implies various means to support discovery and autoconfiguration instead of static prearrangement between parties. Thus, the group is developing proposals to be promoted in various communities, including potential submissions to the OASIS Security Services Technical Committee for consideration as standards.

The main important aspects of their contribution are that the partner keys used to sign and validate SAML SSO messages are included in the SAML metadata document, and trust in these keys is derived from the established trust in the metadata document itself. Also, the metadata document must be signed and the X.509 certificate chain used to validate the signature is included in the document. Thus, each partner just needs to configure the root certificates.

But this idea is not quite different than just relying on X.509 certificates. There are only two ways to establish trust in the metadata signatures: based on metadata signing certificates together with a traditional PKI or using out-of-band certificates as a form of pre-shared keys for signature validation. The proposal is focused on reducing the manual steps but it does not address dynamism in the sense of trust establishment and evolution and does not consider risk assessment. Although the process is lighter and federations are established more rapidly, trust continues to lie in pre-established relationships, with no evolution over time, and entities cannot take autonomous decisions without some preconfigured information. Furthermore, the proposal is tied to certificate-based trust decisions and it is focused on the web SSO profile, but a more general solution is needed that can be applied to a broader range of federation use cases and protocols.

Another related work, carried out by Boursas et al. [Boursas and Danciu, 2008] [Boursas and Hommel, 2006], contributes towards the dynamic management and expansion of Liberty Circles of Trust. The main idea is to maintain a repository where trust relationships are stored. This repository is accesed by the providers in a federation in order to find a path to unknown entities and derive transitive trust relationships. Basically, they define algorithms and work-flows to allow the establishment of dynamic transitive relationships. However, problems such as the exchange of trust information over current federation protocols, are not addressed. In other paper by Boursas et al. [Boursas and Hommel, 2007], they complete the initial proposal and also introduce a notion of risk management. Risk values are used together with trust to define an enhanced access control to service provider resources. In regard to risk quantification, they just assume that resource owners always specify risk levels associated to their resources. Then, based on this assumption they exemplify it by means of a four level linguistic scale (low, medium, high, and critical risk). The final decision of granting access or not is made by requiring higher trust values when the risk increases. It is to note that this approach does not evaluate risk and trust to establish a federation, but to grant user access to service provider resources.

Bertino et al. in [Bertino et al., 2007] propose to introduce Automated Trust Negotiation techniques (ATN) [Skogsrud et al., 2004] in identity management frameworks as a mean to allow dynamic cooperation. Thus, negotiating parties establish trust between them through bilateral credential disclosure. Their approach, called FAMTN (framework-federated attribute management and trust negotiation) supports negotiations between an SPs and the user, and between two SPs in the same federation. Such a negotiation aims to establish a trust level sufficient to release sensitive resources, which can be either data or services.

In other recent work [Kylau et al., 2009], Kylau et al. identify possible trust patterns in identity federation topologies and enumerate a number of risks as the basis to discuss the trust requirements of each pattern. In this regard, Also Jøsang et al. [Jøsang et al., 2005] define trust requirements for several identity management models including the federated case. These works set an important foundation for modeling trust in federation scenarios. Furthermore, a conclusion that can be easily extracted is that current federation specifications assume that a trust relationship exists but there is no standardization about how to create or manage it in an automated and secure manner.

All the above mentioned works preceded our proposal and motivate our ideas, but afterward new research also started to grow around the field. The work by Gómez-Mármol et al. [Gómez Mármol et al., 2010] presents TRIMS, a "*Privacy-Aware Trust and Reputation Model for Identity Management Systems.*" Their model is oriented to web scenarios with three kind of entities: WSPs, WSCs and users. The WSPs are those entities providing identity attributes (i.e., age, e-mail, location, etc.); while the WSCs are the entities actually delivering the requested web service (e.g., a film, a book, etc.) to the users based on the identity information. Thus, when no Service Level Agreement (SLA) exist to handle the exchange of identity information between providers in a secure manner, TRIMS can be applied to calculate trust and make a decision based on the computed trust value. The trust value that a WSP assigns to an WSC is an aggregation of users opinions, other WSPs opinions and history of transactions. The final trust value refers to general trust since it does not define trust and/or reputation contexts. Furthermore, risk is not included and the protocol exchange to gather trust data is not detailed. In [Zuo et al., 2010], Zuo et al. present a solution to achieve dynamic identity federation based on the introduction of ATN techniques. In this sense, there is a similarity with the work done by Bertino et al. [Bertino et al., 2007]. However, [Bertino et al., 2007] applied ATN to improve access control to user attributes while Zuo's proposal applies ATN to negotiate on federation establishment. This differentiation implies that Zuo's ideas are closer to our work. More specifically, the authors propose an architecture together with a new information exchange protocol and a prototype implementation of the dynamic federation framework. The proposal is built on SAML and the trust relationship is only based on punctual negotiation, without monitoring the relationship and the evolution of trust.

Similar research, conducted by Xiang et al. [Xiang et al., 2010] also identifies the need to move from static agreements in federations to a more flexible model. Thus, they propose an underlying network and trust model for dynamic federation, which uses a modified Dijkstra algorithm for the calculation of a trust bootstraping value for unknown parties. Each entity first places an initial trust values for neighbors (i.e., for known entities) based on the existence of digital certificates that verify their identity. Depending if the certificates are issued by a common Certification Authority (CA), by a trusted CA or they are self-signed, the trust level is higher or lower. Then, trust in non-neighbor entities is calculated by modeling the problem as finding the shortest path between the two involved entities and multiplying the trust values for all the links that conform the path. For this purpose, each entity handles a trust table where trust values are stored. The network model consists of federations that have IdP-hubs acting as external interfaces to interoperate with other federations. For the IdP-Hubs to be able to know the network topology and operate over it, they conceptually describe a protocol (DYNFED) similar to link state routing protocols. The protocol is used to announce the local trust levels stored in the entities' trust tables, so IdP-Hubs are able to construct the network graph under their domain. As occurred in all the other approaches, risk is not considered when making decisions. Instead they use a general trust value calculated in a transitive manner.

On the other hand, there are a number of key research projects - primarily funded by the European Commission - that are involved (or have been involved) in identity management related topics. We gather here the most relevant ones. More specifically, within the Seventh Research Framework Programme of the European Union from 2007 to 2013, several new projects related to identity management started. PICOS (Privacy and Identity Management for Community Services) [PICOS Project, 2011] investigates and develops a state-of-the-art platform for providing trust, privacy and identity management in mobile communities. PrimeLife [PrimeLife Project, 2011], that is a continuation of project PRIME [PRIME Project, 2008], works on privacy-enhancing technologies that can enable citizens to execute their legal rights to control personal information in on-line transactions. Thus, the project is advancing the state-of-the-art in the areas of human computer interfaces, configurable policy languages, web service federations, infrastructures and privacy-enhancing cryptography. For this purpose, PrimeLife works with the relevant open source communities and standardization bodies. The SWIFT [SWIFT Project, 2010] project (Secure Widespread Identities for Federated Telecommunications) leverages identity technology as a key to integrate service and transport infrastructures for the benefit of users and the providers. It focuses on extending identity functions and federation to the network while addressing usability and privacy concerns. The research within SWIFT includes a gap analysis to identifie challenges in existing identity frameworks, a requirements list to address these gaps and a generic architecture based on the requirements. As PRIME does, this project has an important activity in standardization organisms. Summarizing, PICOS and PrimeLife are more concerned with privacy whilst SWIFT focuses on improving federation functions. In this sense SWIFT ideas are more related to ours, since they point out the need for easier trust establishment mechanisms and their proposed architecture envisions a module for *"Dynamic Federation and Trust Negotiation"*. Finally, thought all the projects embrace trust and risk considerations to some extent, none of them provides a comprehensive solution to the specific case of deciding whether or not to federate (or transact with a federated party) based on trust and risk calculation, as we develop in this thesis. In conclusion, current proposals do not provide a general solution that address how to extend a federation framework independent of the transport protocol or use case. Finally, we aim to address more dimensions of trust, such as risk management, that are not considered in the presented approaches.

2.4.2. Standards Developing Organizations and Related Bodies

Several standardization developing organizations (SDOs) and related bodies are working on identity management topics that conform fundamental pieces to achieve the establishment of federations in a dynamic manner. In the following, we name this organizations and explain their work and how it relates to the vision of dynamic federation.

Organization for the Advancement of Structured Information Standards (OASIS)

OASIS leads several efforts in the standardization of federation standards. As previously documented in this chapter, SAML, WS-Federation, and Identity Interoperability Mestasystem for Information Cards, are federation frameworks standardized by OASIS. In addition, apart form these mature identity standards,

OASIS created other Technical Comitees (TCs) that are also related to identity management. The most relevant groups that are adrressing issues of trust, reputation and federation establishment are:

- OASIS IDentity in the cloud (IDCloud) TC.

  This group, created in 2010, develops profiles of open standards for identity deployment, provisioning and management in cloud computing. The TC identifies gaps in existing identity management standards and investigates the need for profiles to achieve interoperability within current standards. It performs risk and threat analysis on collected use cases and produces guidelines for mitigating vulnerabilities.

  The most relevant technical work produced by the committee so far are the "*OASIS Identity In The Cloud Use Cases v1.0*" and the "*Identity in the Cloud Gap Analysis Version 1.0*" documents. Both remark that cloud computing is a natural evolution from virtualization and the service provider model, and so it magnifies the need for federating identities between providers and customers. For cloud to succeed, they state, standards must further evolve to make identity federation economical, scalable, and practical for the mass market.

- OASIS Open Reputation Management System (ORMS) TC.

The ORMS group, formed in 2008, has the goal of advancing the ability to use common data formats for representing reputation data. In their main document, "*Open Reputation Data (ORD) Draft Version 0.1*", they describe a reference model for exchanging portable reputation information between reputation data providers and consumers.

- OASIS Electronic Identity Credential Trust Elevation Methods (Trust Elevation) TC.

  The OASIS Trust Elevation TC, created in 2011, works to define a set of standardized protocols that service providers may use to elevate the trust in an electronic identity credential presented to them for authentication. The Trust Elevation TC promotes interoperability among multiple identity providers–and among multiple identity federations and frameworks–by facilitating clear communication about common and comparable operations to present, evaluate and apply identity to sets of declared authorization levels.

European Telecommunications Standards Institute (ETSI)

In 2011 the ETSI Industry Specification Group on Identity and Access Management for Networks and Services (ISG INS) published a specification entitled "*Identity and access management for Networks and Services; Dynamic federation negotiation and trust management in IdM systems*" [ETSI, 2011]. The document describes a problem statement to federation establishment based on dynamic SLA negotiations, (so called "ad-hoc federations"), and presents related use-cases and requirements.

The ETSI specification recognizes that using bilateral static agreements is not feasible for a global scale federated internet. Current procedures to establish federations impose a high barrier to small companies (or even individuals) that act as Service Providers, since they cannot afford the time and money to fulfill those agreements. On this basis, the specification states that techniques are required which give the possibility of ad-hoc federation; and that such techniques must consider reputation, quality of credentials and risk.

The document presents several high level use-cases where dynamic federation is useful and also contemplate the notion of trust establishment based on reputation to allow dynamic

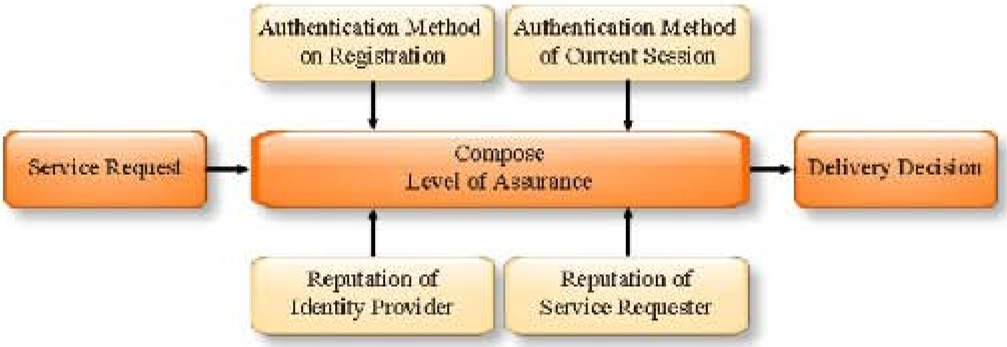

Authentication Method on Registration
Authentication Method of Current Session
Service Request
Compose Level of Assurance
Delivery Decision
Reputation of Identity Provider
Reputation of Service Requester

Figure 2.16: Conceptual scheme for the composition of a Level of Assurance (LoA) metric (©[ETSI, 2011])

interactions.

Basically, the recommendation to achieve ad-hoc federation establishment is based on the image in Figure 2.16. Upon a service request, the Service Provider should calculate a Level of Assurance (LoA) based on four elements, namely: (1) the authentication method used when the user registered at the Identity Provider(e.g., PostIdent, E-Mail verification, etc.); (2) the authentication method of the current session (e.g., username/password, SIM-Card, etc.); (3) the reputation of the user/service requester; and (4) the reputation of the Identity Provider. Thus, the service delivery decision will depend on the LoA and internal risk taking factors.

Finally, it is worth noting that the specification remarks the need for defining metrics to qualify the LoA and the internal risk factors. The document concludes that "*the solutions or mechanisms which will allow to instantiate an ad-hoc federation are still open and should be discussed*".

ABFAB Internet Engineering Task Force Working Group

The Internet Engineering Task Force (IETF) started a working group focused on "*Application Bridging for Federated Access Beyond web*" (ABFAB) [Howlett et al., 2012]. This group envisions federated identity as a mean for facilitating the controlled sharing of information about principals, commonly across organizational boundaries. This avoids redundant registration of principals who operate in multiple domains, reducing administrative overheads and improving usability while addressing privacy-related concerns and regulatory and statutory requirements of some jurisdictions.

They noted that federation mechanisms are in use for the web but not for other contexts. Based on this problem statement, the working group has a special focus on specifying a federated identity mechanism for use by other Internet protocols not based on HTML/HTTP, such as for instance IMAP (Internet Message Access Protocol) or SSH (Secure Shell).

The ABFAB working group is currently working on a series of drafts[10], being the most remarkable ones:

- "*Application Bridging for Federated Access Beyond Web (ABFAB) Architecture*": defines an architecture that addresses the problem of federated access management to primarily non-web-based services, in a manner that will scale to large numbers of identity providers, relying parties, and federations.

- "*Application Bridging for Federated Access Beyond web (ABFAB) Use Cases*": enumerates a list of use-cases describing how technologies based on the the ABFAB architecture and specifications could be used to achieve identity federation in non-web scenarios, such as cloud computing, grid computing, high performance computing, etc.

- "*Application Bridging for Federated Access Beyond web (ABFAB) Usability and User Interface Considerations*": This document provides recommendations to design consistent interfaces for managing user's identities.

Regarding trust establishment and the creation of dynamic relationships between providers to share user identity data, the group points out the need for a "*Trust Router Protocol*". A Trust Router Protocol, they state, allows a new partner to be added to an ABFAB community by peering with any member of the Trust Router network, instead of requiring configuration changes by every partner who may wish to connect with the new partner. Thus, its main function is the distribution of information about existing trust relationships within the partnership, avoiding the operational costs and limitations of using a Public Key Infrastructure (PKI).

[10] http://datatracker.ietf.org/wg/abfab/

Kantara Initiative

Kantara Initiative was announced on 2009, by leaders of several foundations and associations working on various aspects of digital identity. It is intended to be a robust and well-funded focal point for collaboration between members of the identity community. As stated earlier in this chapter, Kantara is not a standards body but submits recommendations to standards bodies.

The organization is structured into working groups that deal with different aspects of identity management. The groups that are more related to the concept of dynamic federation are:

- *Trust Framework Meta Model Work Group.*

  This group works on defining the components of a Trust Framework and providing a mechanism for comparing Trust Frameworks developed by communities. The Trust Framework Meta Model would be a reference resource not only within Kantara Initiative activities but also for any community seeking to understand the Trust Framework concept and potentially as guidance toward the development of Trust Framework components.

  It is worth noting that their work in progress considers a federation use-case that builds on the notion of reputation-based trust. Although it is not yet defined, they state that "*reputations-based systems have a process that will compute trust based on behavior of claimants and rating of members. This is not necessarily limited to humans, but could be used for IdPs and other roles in federation as well.*"

- *Federation Interoperability Work Group.*

  The purpose of this group is to profile existing specifications to define an interoperable trust infrastructures for use by parties participating in Trust Frameworks. This will allow entities to determine the certification status and configuration parameters of entities outside of their local federation.

- *Business Cases for Trusted Federations Discussion Group.*

  The purpose of this discussion group is to identify and raise awareness of business cases around the deployment and adoption of federation models and systems Ű

> particularly the Trust Framework model.
>
> This group will gather input from International stakeholders specifically, actors from within vertical and jurisdictional communities of trust with the purpose of allowing participants to share information about successful and challenging experiences with specific focus on the business drivers and motivations for deploying federations and the Trust Framework model.

Finally, among the work developed in Kantara, it is worth to mention their certification programs. Kantara Initiative has answers to both technical and operational assurance needs through certification programs designed to give the marketplace confidence, consistency and control when deploying identity solutions. More specifically, the programs are: (1) *Identity Assurance Certification Program*, that certifies identity credential systems based on four distinct Levels of Assurance ; and (2) *Interoperability Certification Program*, that tests identity management implementations and certifies the degree of fulfillment of the standards.

The Open Identity Exchange (OIX)

The Open Identity Exchange (OIX) [OIX, 2013] is a non-profit organization dedicated to building trust in the exchange of online identity credentials across public and private sectors. OIX also received initial grants from the OpenID Foundation (OIDF) and Information Card Foundation (ICF) to advance assurance for open identity technologies. The initial members are Google, PayPal, Equifax, VeriSign, Verizon, CA and Booz Hamilton.

The OIX proposes an Open Identity Trust Framework (OITF) as a solution to enable large scale networks of trust, facilitating cooperation of IdPs and SPs without the requirement of an agreement. An OITF is defined as a set of technical, operational, and legal requirements and enforcement mechanisms for parties exchanging identity information. In an OITF additional actors look after these requirements and mechanisms to support the flow of information among users, IdPs and SPs. The roles and relationships of these additional actors are shown in Fig. 2.17.

Policymakers start by deciding the technical, operational, and legal requirements for exchanges of identity information that fall under their authority. They then select *OITF Providers* to implement these requirements. These OITF Providers translate the require-

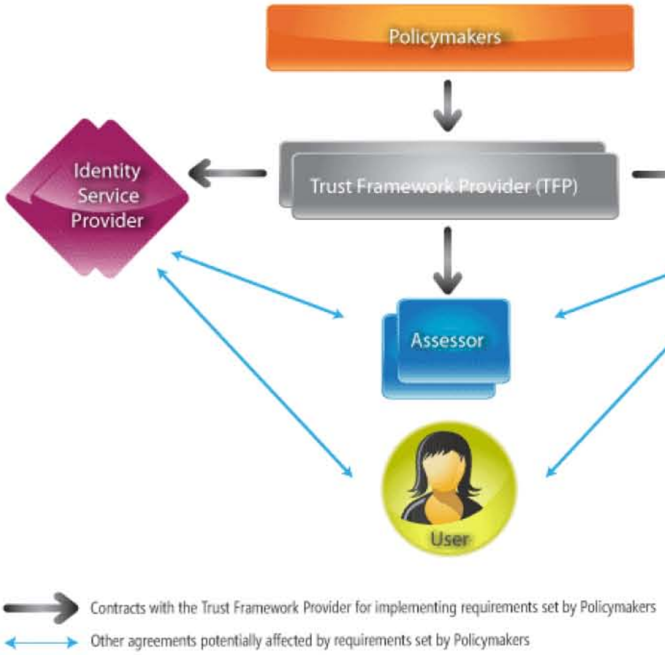

Policymakers
Identity Service Provider
Trust Framework Provider (TFP)
Relying Party
Assessor
User
Contracts with the Trust Framework Provider for implementing requirements set by Policymakers
Other agreements potentially affected by requirements set by Policymakers

Figure 2.17: Roles and relationships between the participants in an OITF (©[OIX, 2013])

ments into a blueprint for a trust framework. *Assessors* evaluate identity service providers and relying parties and certify that they are capable of following the OITF Provider's blueprint. The OITF Provider vets identity service providers and relying parties and contracts with them to follow its trust framework requirements when conducting exchanges of identity information. The contracts carry provisions relating to *dispute resolvers* and *auditors* for contract interpretation and enforcement. Requirements flow down through agreements, as shown in the directional arrows in Fig. 2.17.

European Network and Information Security Agency (ENISA)

European Network and information Security Agency (ENISA) is as a body of expertise, set up by the EU to carry out very specific technical, scientific tasks in the field of Information Security, working as a "European Community Agency". It is not a standards developing organization but its work includes reviewing of standards from the security point of view and providing best practice recommendations that are useful for SDOs.

ENISA regularly publishes documents in regard with identity management, where the need for trust establishment is always contemplated. In "*Managing Multiple Electronic*

Identities", trust is recognized as a central issue in all identity transactions. The document states that trust must be established between the different actors involved in identity federation. The subject must be confident that their personal information will be handled appropriately, in order to take part in the transaction. Equally, the relying party must be confident that the subject's obligations under the transaction (such as payment) will be honoured. In order for the relying party to have this confidence, they will need the means to assess the trustworthiness of the assertion being provided in that particular context. However, the document concludes, at present there is no reliable mechanism underlying this process.

It is also remarkable the work of the agency in the field of risk. The ENISA working group on risk management regularly publishes a variety of information pertinent to Risk Management and Risk Assessment.

More specifically, their work includes inventories of methods tools and good practices, achieved results in the area of emerging eisks information material for Small and Medium Enterprises (SMEs), comparability and interoperability issues of methods, tools and good practices Integration issues of Risk Management with other operational processes.

Chapter 3

Understanding the Problem of Dynamic Federation

A problem well stated is a problem half solved

Inventor Charls Franklin Kettering (1876-1958)

Contents

3.1. Chapter Overview

Defining a research problem is the fuel that drives the scientific process, and is the foundation of any research contribution. But the process of choosing the research topic and carefully stating the problem to be solved defining all its nuances is not a trivial task. In fact, a lot of unease questions arise: *What?*(What are the major questions for the topic?), *Where?* (Where is the topic important: at the local, national or international level?), *Who?* (Who are the information providers on this topic? Who is affected?), and the ubiquitous and most important question... ***Why?***.

There are also frequently referred points[1] to consider in finding and developing a research topic: that is compelling and interesting, that is original, that is worthy, that is a solvable and manageable problem, that leads to more research questions, etc. Between them, there is also a more personal (but not for this less important) point: to choose a problem that can be enthusiastically pursued.

Given the importance of having a well-defined problem to understand the proposed solution, we dedicate this whole chapter to this endeavor before starting with the development of the thesis. Our aim here is to demarcate the problem area considering the above mentioned questions and points, to illustrate what caused the need to do this research and to provide the reader with a clear understanding of the research problem we address. We want to show how we found a gap in knowledge, and how we are seeking to fill it.

But above all, our intention here is to transmit **why** the chosen topic kept us motivated during years and still ignites enthusiasm to derive exciting new lines of research; to convey the reasons that drove us in the quest to fill the particular gap in knowledge that will make Dynamic Identity Federation a reality.

Although the thesis purpose and objectives were already precisely declared in the Introduction (Chapter 1), it is only after the revision of the state-of-the-art and related work that we can provide a deeper explanation. Thus, the following sections focus on explaining the problem area; analyzing the existing gaps, including how they are being tackled and which questions remain open; and providing an overview of the contributions of this thesis to fill existing gaps. Through all of these sections, the factors that motivate our research are highlighted.

[1] e.g., the chemistry professor and author Robert Smith, in his book *"Graduate Research: A Guide for Students in the Sciences"* (ISI Press, 1984), lists 11 points to consider in finding and developing a research topic

3.2. Federation and the Identity Landscape

"The world is already federated, it's the computer that needs to catch up, specifically the security protocols."
-Gunnar Peterson, 2010[2]

Identity is a crucial element in most computer security mechanisms and currently located at the core of the internet economy. The way identity has been handled online evolved following the progression of Internet technology closely. The famous cartoon stating [3] "*On the Internet, nobody knows you're a dog*" illustrated soon the concerns about privacy and authentication posed by online identity management. During the formative years of the web, when computers were not hyper-connected, password-based authentication was the first approach to handle identity. This mechanism worked pretty well at that time, due largely to how little data they actually needed to protect. A user password was limited to few sites, such as an email account and maybe an e-commerce site or two. However, with the advent of web services, the explosion of online applications and the increase of online transactions, identity got far more complex. Today, we are asked to prove our identities every time we board a plane, check into a hotel, make a purchase via check or credit card, or log onto a computer or secure web site. Many large scale studies point out the high number of accounts a typical web user has (around 25 as found in [Florencio and Herley, 2007]). Users face the burden of managing this increasing number of accounts and passwords, which frequently leads them to devise password management strategies that degrade the security of their protected information [Gaw and Felten, 2006]. In the enterprise world, centralized IdM solutions were created to deal with users and data security where the user and the systems accessed by the user were within the same network.

Thus, identity management has been historically implemented following a *"silo"* model. Each service provider or organization creates and maintains the identity management process, incorporating means for identifying, proofing, provisioning, authenticating, securing, managing and otherwise maintaining the base of users. But these identity schemes are inefficient for the current demands of cross-organization cooperation, partnership and collaboration. PKI [Adams, C. and Farrell, S., 1999] infrastructures were envisioned once as a universal solution for the identity problem. Nevertheless, though the mechanisms associated with X.509 digital certificates worked well for identifying computer systems and

[2] http://1raindrop.typepad.com/1_raindrop/federation/

[3] Steiner, P. (1993). On the Internet Nobody Knows You Are a Dog, New Yorker (69)20, 5 July.

establishing secure communication channels, global PKI never happened as an identity management system.

Federated Identity Management or FIM is the newest approach to distributed identity management. It frequently relies on PKI for the distribution of cryptographic material, inheriting partly the complexity burden of these infrastructures. However, federation loosens the requirement for a single root authority by replacing the root with definitions, policies, and semantics agreed upon out of band, and provides new identity functionalities. A federated identity is a single user identity that can be used to access a group of web sites bound by the ties of federation, which reduces the burden of having to manage different credentials for every site. The ultimate goal of identity federation is thus to enable users of one domain to securely access data or systems of another domain seamlessly, without requiring redundant user administration. The value of federated identity management is its simplification and decoupling of functions: authentication is separated from the process of accessing resources. According to this scheme there are three roles (already detailed in Chapter 2), namely the IdP, the SP and the user. And everyone can benefit from the federated model. The IdP can focus on improving the process of authentication, perhaps providing different modes of strengths of authentication, perhaps providing other services. The SP no longer has to handle authentication - a messy, problematic business - and can focus instead on the provision of services. And the user only has to log in once with a single set of credentials. Furthermore, federation is not only about authentication reuse, but it allows also the sharing of identity attributes and authorization information, which leads to other use-cases such as cross-domain user account provisioning, cross-domain entitlement management and cross-domain user attribute exchange The Kantara Initiative, actively involved in the advance of online identity management, highlights four basic value propositions for FIM. Namely:

- Economics of scale: The reuse of credentials shares the cost of provisioning and supporting credentials and their use.
- Information Security and privacy: The reuse of credentials makes it feasible to have strong credentials and procedures in place to protect identity-related security objectives. As a result it is possible to make processes available on-line where the risk would have been too high without strong protection in place. Another benefit is to reduce identity-related fraud.

- Business enabling: Whereas applications with a high value per transaction or per user usually can afford their isolated IdM, long-tail [Anderson, 2004] applications cannot. The availability of better price/performance credentials will enable new applications.
- Improving existing processes: Many on-line applications loose prospective customers or users when they require registration. Confirmation mails lost in SPAM-filters are a frequent cause. Instant identification of a user would result in a higher ratio of users continuing through the full process.

Impulsed by all these benefits, the FIM model entered the identity landscape around a decade ago and is becoming more and more a hot topic. But federation is easier said than done. While some important federation efforts have repeatedly failed, such as the Microsoft's flagship identity solution Cardspace, the FIM experience of the last decade was not entirely negative. Successful FIM deployments have been conducted in education and industry [Landau and Moore, 2011] mainly based on SAML, a very alive standard that announced a newer version in December 2012. Also, in the social web arena, OpenID and OAuth are working well.

However, the wide scale deployment of federation still did not happen. Between the barriers that are hindering the adoption of federation at a global internet scale, we think that the underlying trust models are the root cause. In the next section, we take a closer look at the existing gaps in the current federation schemes in order to clarify the reasons that led us to propose a **dynamic federation** model.

3.3. Gap Analysis

A gap, according to the definitions in [Singh, 2005] is the lack of (or lack of adoption of) a solution based on open standards or specifications to support a specific industry need or requirement, the lack of a specific feature, or an incomplete capability.

A gap, they say, can arise from the lack of a technical mechanism or protocol, a best practice or guidelines specification, or a performance specification. It can also arise from the lack of a specification describing the application of a defined technology to address specific network architectures (e.g., NGN and IMS), business models, and assumptions (e.g., scalability). Or it can arise from the lack of a sufficient administrative mechanism or national mandate.

Considering the above definitions and based on the literature review in Chapter 2 plus the available studies about barriers for federation adoption [ITU-T Focus Group on Identity Management (FG IdM), 2007] [Landau and Moore, 2011] [Jensen, 2012] [Jøsang et al., 2005] [Sun et al., 2010], the main gaps (non-exhaustive) found in the FIM field are enumerated below. Jensen in [Jensen, 2012] does a good work in compiling the challenges, so we fundamentally base on his text:

1. **Trust.** Trust is the fundamental concept underlying federations and is not surprising that challenges related to this concept are the most frequently raised in the literature. The separation of functions in the federated model implies that each involved party has to rely on the other: e.g., the SP has to trust the IdP to correctly authenticate the users that will access its services; and the IdP has to trust the SP to properly handle the user identity data. A trust relationship must thus exist. Currently, trust is established based on static agreements that must be set by administrators out of band before interaction, which makes the federation procedure slow. In other cases, trust is assumed to exist by default. Pre-established trust relationships limit the cooperation with potential new business partners, while not applying a trust model poses important security risks. Making the federation process more agile is specially relevant for dynamic and short-term collaborations, where the time and cost of setting up a federation in the traditional way will probably not outbalance the rewards of cooperation (see Section 3.4.3 for potential use-cases). For these reasons, current specifications come short in presenting solutions for dynamic environments, and there is an important gap here.

2. **Security.** A serious concern in FIM is identity theft. Impersonation attacks with stolen credentials are in fact very common nowadays. In the FIM case, problems regarding a stolen identity affect all federation partners. An impersonation attack can be performed by stealing an authenticated user's security token, and then use it to access resources in the federated environment [Han et al., 2010]. A consequence of this is that even systems with more secure authentication protocols than username/password schemes, such as systems using two factor authentication, are exposed to identity theft threats.

There are important works on federation security research. Gro*β* et al. [Groß, 2003] initiated the security analysis of SAML-based SSO finding deficiencies in the information flow between the involved entities, which influenced a revision of the standard. More recently, the paper *"On breaking SAML: be whoever you want to be"* by Somorovsky et al. [Somorovsky et al., 2012] showed how integrity protection in SAML can be successfully circumvented by application of different XML Signature specific attacks, so attackers can take *"whatever identity they want"*.

In general, FIM specifications contain security recommendations and a variable number of configurable security features. Depending on the selected options in implementation the achieved security level will be different.

3. **Privacy.** Privacy is a challenge in FIM [Glässer and Vajihollahi, 2010], and a recurring topic among researchers. The key goal in FIM is to share personally identifiable data, while at the same time guaranteeing privacy, i.e., taking into account issues such as data protection, user consent and compliance with legislation. In the case of FIM systems, due to the distributed sharing of identity information, the security domain is blurred, which causes extra challenges. The enforcement of privacy policies is a big concern, as pointed in [Squicciarini et al., 2008]. Even though there exists privacy policies, and users express their privacy preferences, there are no requirements to enforce these policies and preferences through technology. Users should be able to regulate the release and use of their own identity information, but specifications do not properly cover user empowerment in the control of their identities. There is a lack of support to match users' consent with privacy policies and to configure attribute release policies (ARPs) . Only Shibboleth v1.2, a SAML-based solution, supports simple ARPs, but in a proprietary format.

4. **Interoperability.** Despite the standardization efforts in the FIM domain there are still considerable challenges related to interoperability. As we detailed in Chapter 2, there is a number of different specifications for identity federation. FIM protocols work well in homogeneous environments where all collaborating partners use the same standard. However, there are situations, where partners adhere to different standards, and this increases the complexity. Mechanisms to bridge protocols, to perform token conversions and to combine FIM technologies are being developed to address this gap [Pacyna et al., 2009] [Monjas et al., 2009]. But even with the use of

one standard, there might be also interoperability challenges. In the case of SAML, the high number of available protocol options and conformance variations are aspects that cause difficulties. Furthermore, the process of determining what are the necessary identity attributes, and finding common data schemes for interorganizational cooperation may be challenging [Bertino et al., 2009].

5. **Assurance.** According to Baldwin et al. [Baldwin et al., 2007], identity assurance is the process of ensuring that identity management is under appropriate control. Existing FIM frameworks do not provide information about the verification of identity data of the individuals enrolled and stored at each IdP. In FIM the identity management process involves different organizations, which may have different risk policies. The assurance process, i.e., demonstrating that controls and processes are being followed and sufficient to mitigate identified risks, is vital [Baldwin et al., 2007]. Without properly addressing assurance considerations, it is unlikely that FIM systems will be adopted for enterprise tasks. Currently, the SAML specifications are evolving to incorporate and convey information regarding the level of assurance.

6. **Liability.** Liability is the state of being legally obliged and responsible. It is widely believed that the inability to solve the liability issue is at the root of the slow adoption of federation technologies [Camp, 2010]. There are also challenges derived from the fact that federated systems may span across multiple jurisdictions. In these cases, there are differences especially at national level related to privacy, labour, and disclosure laws. All potential liability issues must be addressed before federating identity and access systems, indicating what will happen and who is responsible if something related to the FIM process goes wrong.

7. **IdP Discovery.** IdP discovery is the process of determining where to send authentication requests when a user wants to access an identity based service [Maler and Reed, 2008]. This is a problem in service-initiated Single Sign-On use cases, since service providers (SP) can be configured to accept security tokens from numerous identity providers. A common solution to this problem is to directly ask the user, providing him with a list of IdPs from which he has to select the correct. This is especially a problem when the list of possible IdPs gets extensive, affecting also usability. Another more elegant solution is to give the users a smart client that is smart enough to know the answer. This is a compelling challenge and specifications are neutral

about the discovery process. Only OpenID provides a discovery protocol, but also present challenges for hosted domains. The OpenID Foundation is currently working to create a next-generation OpenID discovery protocol that address the open issues.

8. **Bussiness model.** A big issue in the adoption of federation, as for every other technology, is business. As pointed in the recent paper [Landau and Moore, 2011], the benefits for all the involved parties must be balanced for a federation to succeed. The study in [Sun et al., 2010] claims that the wide adoption of federation is not happening because service providers do not have sufficient incentives to become relying parties. There is a need to define models for monetizing identity services and investigate business needs.

Of all the above challenges, we aim to address the trust establishment issue. The trust topic has been widely investigated in computer science, where research contributed to the development of modern open distributed and decentralized systems. Trust has been studied in the context of decentralized access control, public key certification, and reputation systems for P2P networks. Now, with the emergence of distributed identity management, it is important to evolve research in this context.

The specific research question we aim to solve is ***"Can trust in FIM be established in a more dynamic yet secure way?"***. These considerations on dynamic trust aspects assume an open federation environment (as Internet), where entities may come and go dynamically, and thus may not be relevant in closed environments with more static participants. But is our aim to contribute in this specific gap, since we are convinced that it will foster new ways of collaboration and use-cases that will have an important impact on the way business are conducted on the Internet (see Section 3.4.3).

On the other hand, our approach considers security, privacy, legal aspects and assurance as basis for measuring the risk in cooperating with another entity. The combination of risk and trust will be used to make the final decision on cooperation.

Finally, before delineating our proposal and further clarifying its potential impact, we summarize in Table 3.1 how current related work covers trust establishment issues. The main goal with this is to highlight the novelty and differences of the ideas presented in this thesis.

Proposal	Year	Approach	Algorithms	Dissemination	New Elements
[Bertino et al., 2007] (FAMTN)	2007	ATN based	Policy Negotiation	Policy, Tokens	Trust Logic, Policy Negotiation Protocol
[Boursas and Danciu, 2008]	2008	Trust and Reputation (graph based)	Trust path: Breadth-first search, Trust and Reputation Computation: hand-crafted	Central LDAP directory with trust matrix, Protocol primitives not defined	Trust Logic, Dissemination Protocol
[Zuo et al., 2010] (DFed)	2010	ATN based	Policy Negotiation	Policy Negotiation Protocol (over WS-Federation)	Trust Logic, Dissemination Protocol
[Xiang et al., 2010]	2010	Trust and Reputation (graph based)	Trust path: Dijkstra, Trust Computation: hand-crafted, Reputation: PageRank [Page et al., 1999]	DYNFED protocol: based on link state routing protocols, primitives not defined	Trust Logic, IdP Hub (pre-configured), DYNFED protocol
[Gómez Mármol et al., 2010] (TRIMS)	2010	Trust and Reputation (storage at IdPs)	Trust and Reputation Computation: hand-crafted, Homomorphic Encryption for privacy	Global Storage at IdPs, Protocol primitives not defined	Trust Logic, Dissemination Protocol
Dynamic Federation	2009-2012	Trust, Reputation and Risk (distributed)	Trust and Reputation Computation: pluggable, Risk Computation: metric-based, Aggregation: Fuzzy-based	IdMRep Protocol (over SAML)	Trust Logic, Dissemination Protocol

Table 3.1: Summary of related work compared to our proposal

The proposals shown in Table 3.1 are the most relevant and closer to our ideas of those presented in the related work. For each of the proposals, we show the year of publication, the type of approach, the algorithms used, and the dissemination mechanism to convey trust data. There are two kinds of approaches, namely: based on ATN and based on trust and reputation. Our approach (last row in the table) falls into the second category and the principal novelty introduced is the risk model. Decisions are made on the basis of trust reputation and risk, while on the other cases only trust is taken into account. Furthermore, the solution is flexible and not tied to a specific trust evolution algorithm. Instead, any function may be plugged in our trust model as desired by the implementers.

Finally, from the central research question we propounded, a number of subquestions arise: What information elements are most suitable for deriving measures of trust, reputation and risk in FIM? How can these information elements be captured and collected? What principles should we use for designing such systems? How should this information be combined into the decision process? What role can these systems play in the business model of commercial companies? Do these systems truly improve the quality of online trade and interactions?

3.4. Towards Dynamic Federation

3.4.1. Previous Clarifications

With the goal to move from the conventional static bilateral agreements to automated dynamic federation, we propose a federation scheme based on the combination of trust, reputation and risk that allows trust relationships to be established on-demand. Here we clarify the main assumptions and definitions in which our work relies.

Firstly, there are some subtleties inherently related to the term federation in identity management scenarios that must be first clarified to properly understand the proposed taxonomy. So, the verb federate as defined in the Merriam Webster Dictionary[4] means: *"To link or bind two or more entities together"*. In the context of our research there are two possible senses in which the word federation (and variants) can be employed and still being coherent with the former definition. According to it, we can talk about: (a) federation, meaning the act of establishing a relationship between providers; and (b) identity federation, which exists when there is an agreement between various providers on a set of identifiers and/or attributes that are used to refer to a subject (i.e., the user). Thus, an identity federation is not possible if providers are not federated.

But provider federation depends on the identity framework being used: it could be a trivial process, as occurs with OpenID, where no trust model is required to cooperate; or a extremely complex task, as happens with SAML-based systems, where contractual frameworks must be statically established to set up a Circle of Trust.

We aim at designing an infrastructure flexible enough to be applied to the different existing FIM protocols. However, whenever particularization is required we will base our framework on SAML, since this is the most consolidated and complete specification.

During this document we will take into consideration the well-known FIM baseline constellation of 3 actors (User, Identity Provider and service Provider) depicted on Figure 3.1,

[4] http://www.merriam-webster.com/

where a user access a service in the SP based on the authentication performed by the IdP.

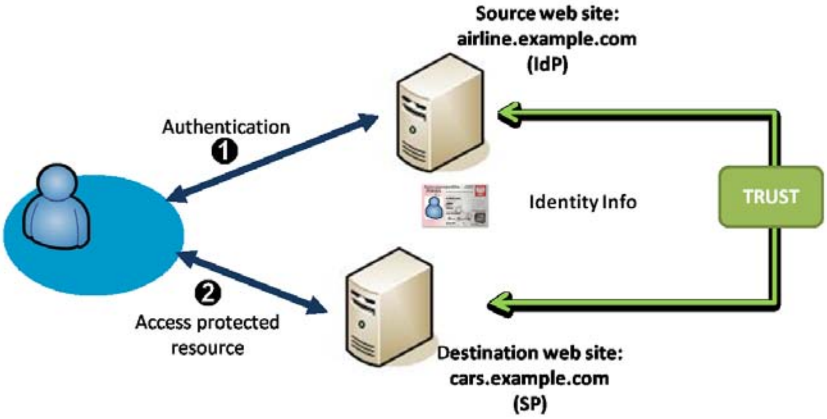


Figure 3.1: Traditional triangle of parties involved in an identity information exchange.

We assume that the trust relationship between the providers may not exist previous to interaction. In such a case, our infrastructure allows providers to decide: a) whether to federate or not with the other and to which extent; and b) once federated, whether to perform each particular transaction or not.

Furthermore, apart from these two kind of decisions, another use-case is possible: selection. For example, if a SP has to select an IdP between a set of available providers, our infrastructure permits to rank all the possibilities and so the best option can be selected.

The quantification of trust and risk and the monitoring of their evolution make possible to have more granularity in the decisions. Thus, the initial set of permissions and types of allowed transactions assigned to an entity can change according to the variation of trust and risk. Though not shown in the picture, the SP and the IdP will probably have relationships with a number of other providers, and these relationships are used in our framework to acquire trust knowledge.

Next, we illustrate how the concepts of trust and risk are treated in this thesis.

3.4.2. Our Vision on Trust and Risk Relationship

We understand both trust and risk as complex multidimensional concepts. In the case of trust, reputation is treated as one of its dimensions. The other dimension of trust we consider is authentication. With these two components it is possible to determine trust: we know about the entity's behavior, and we know about its identity (whether the entity is who it claims to be).

But, as stated in [Solhaug et al., 2007] [Jøsang and Presti, 2004], we believe that trust is not always enough as a basis for the decision to cooperate, but a notion of risk is also necessary. Risk is dependent on both the context of the application and on the assets in place. In the FIM case, we have identified a number of dimensions that contribute to the overall risk in making a decision (see Chapter 5).

Regarding the relationship between both concepts, trust and risk, we agree with the mediating relationship model: the existence of trust reduces the perception of risk, which in turn increases the willingness to transact. Furthermore, as pointed out in [Solhaug et al., 2007], the relationship is affected in extreme cases. For example, there are risks that are too severe that will put a whole enterprise out of business, so they cannot be accepted even if the trust level is good. Similarly, for untrusted entities or those with a very low trust value, the risk should not be accepted even if is not too high.

Regarding the formation of trust and risk values, we agree with the point in [Gefen et al., 2003] that while some antecedents of trust and risk may be the same, there are others, which are not. For instance, a legal contract between two parties enables a transaction by reducing risk, but does not affect trust. More specifically, Figure 3.2 conceptually show how trust and risk are constructed and combined in our model.

It is to say that we use two different terms: *trustworthiness* and *decision trust*. Trustworthiness refers to the trust value that is placed on the entity under evaluation, whereas decision trust refers to the final value obtained after combining trustworthiness with risk.

As shown in Figure 3.2, trust is derived from authentication and behavior data; whereas risk is calculated from policies, SLAs, metadata and history of interactions. The common antecedent in the formation of both values is precisely this history of interactions. Finally, the aggregation of risk and trustworthiness is performed based on a set of rules that model

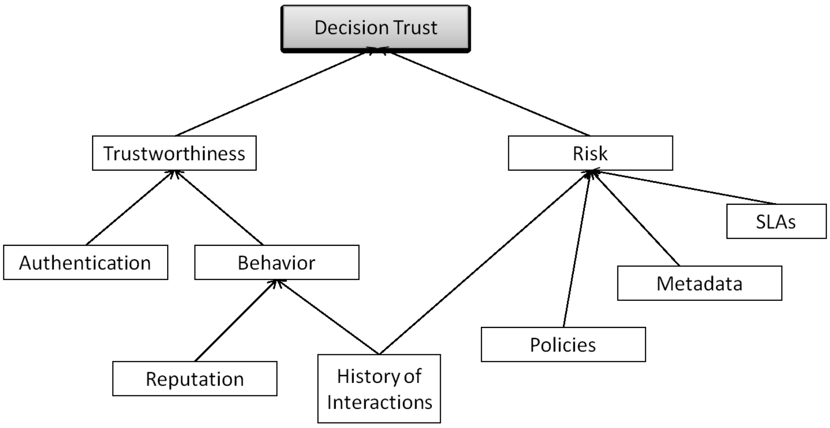


Figure 3.2: Trust and risk formation.

the mediating relationship and take into account extreme values (see Chapter 7).

3.4.3. Potential Impact

Apart from the clarification of the problem we are solving and the assumptions we are making, it is equally important to provide a description of the potential impact that our solution may have. Here we elaborate our views on this issue.

Though business models should be properly investigated for FIM to be deployed at a wide scale, we aim to highlight three potential scenarios where dynamic federation fits and can provide enhancements:

- **Dynamic Business Networks.** This type of networks, also referred as Smart Business Networks, are defined as IT-enabled platforms for dynamically linking different businesses having different 'capabilities' to build a 'networked business enterprise' with innovative business strategies for competing in the changing markets and environmental conditions [Vervest and Zheng, 2009]. The main concept in this business paradigm is the formation of virtual organizations (VOs) or virtual enterprises (VEs). Even more dynamic are the so-called instant virtual enterprises (IVEs), - temporary business entities executing dynamically composed, global business processes to achieve a specified business goal [Grefen et al., 2009].

  VO/IVE creation is opportunity driven. The possibility of rapidly finding a set of partners that best fit a concrete business opportunity and quickly configure them into a collaborative network to exploit that opportunity seems indeed a desirable scenario to face the challenges of market turbulence.

In this kind of organizations, the rapid federation of the identity management systems is required. Users of the different organization parts would probably have to access resources in the domain of partners, and the creation of duplicated user databases at each partner is not a good solution. Identity federation would allow a seamless user experience across all the involved partners without the need of additional efforts in IdM infrastructure. But today's FIM technology does not satisfy high demands of agility. Thus, dynamic federation is key to enabling the complex ecosystems of dynamic business networks.

The same idea of dynamic virtual organizations is also very appealing in other non-business oriented contexts. An extreme case being the incident management and disaster rescuing processes, when it is necessary to very rapidly engage and coordinate activities of a large number of entities (e.g., fire brigades, police, hospitals, local government, non-governmental organizations). This very idea of groups of organizations being able to rapidly configure themselves into some form of mission/goal-oriented collaborative form embeds the notion of great agility [Camarinha-Matos et al., 2005]. Regarding FIM application, the usage of SAML in these scenarios is being investigated [Tran and Wietfeld, 2009].

- **Cloud Computing.** Cloud computing is changing the way industries and enterprises do their businesses in that dynamically scalable and virtualized resources are provided as a service over the Internet. As an example, an enterprise that has a private cloud may want to burst workloads to a public cloud vendor. Enterprise users will end accessing many applications on hybrid cloud computing environments, which go beyond the boundary of the enterprise data center. In these scenarios, Single Sign-on is a challenge. The enterprise typically uses access management to integrate applications in different domains to an application portal, so that the end user can access applications without re-authentication. Access management might work well for the applications within the data center or within the same domain. However, the cloud computing services are typically external to the data center and located within a different domain and shared with multiple other tenants. Federation is thus a useful technology in cloud computing. In fact, the integration of FIM in cloud has been signaled as necessary in many recent research works and technical publications [Gopalakrishnan, 2009] [Sengupta et al., 2011].

Furthermore, elastic cloud computing envisions that cloud services are to be contracted on demand depending on the current needs of each enterprise. In this sense, dynamic federation becomes an enabler to provide agility.

- **Consumer Electronics.** The continuous advance in consumer electronics leads to new scenarios (e.g. digital TV, media distribution, etc.), which evolve and offer new experiences and interactions to consumers with a multitude of providers; highlighting the importance of the role that mobile devices play in such environments. So, identity management is required in order to: a) avoid users dependence on fixed infrastructures, b) allow the interoperability between separated domains, and c) support users' mobility, content adaptation and sharing, services delegation, device heterogeneity, among others.

Summarizing, among other applications, dynamic federation will lower barriers for plug and play B2B integration. It will permit to take dynamic decision on federation by calculating trust and risk on a per-transaction basis. Furthermore the risk evaluation mechanism can be also used alone to asses the own risk level that the particular FIM configuration of the entity poses. Finally the quantification of trust and risk allows also to rank a set of providers that are potential collaborators and decide what is the more appropriate.

3.5. Conclusions

Identity management systems cannot be centralized anymore. Nowadays, users have multiple accounts, profiles and personal data distributed throughout the web and hosted by different providers. However, the online world is currently divided into identity silos forcing users to deal with repetitive authentication and registration processes and hindering a faster development of large scale e-business. Federation has been proposed as a technology to bridge different trust domains, allowing user identity information to be shared in order to improve usability. But the reality is that FIM has not been broadly deployed in the wider Internet. It has functioned well in sectors in which the parties had first established contracts. On the “open” Internet, where IdPs and SPs might not previously have had a relationship, the uptake on federated identity management has been very slow. Consequently, further research is required to shift from the current static model, where manual bilateral agreements must be pre-configured to enable cooperation between unknown parties, to a more dynamic one, where trust relationships are established on demand in a fully automated fashion. We aim to address this gap by introducing a dynamic federation model based on the combination of trust and risk assessment, which are computed whenever a decision of potential collaboration between entities has to be made. We have responded the initial questions suggested in the overview of the chapter, and also stated the main research question. Next, we concentrate on the remaining interrogation: ***How***?

Chapter 4

Architecture Proposal

"As a maturing discipline with no clear rules on the right way to build a system, designing software architecture is still a mix of art and science."

-SoftwareArchitectures.com

Contents

4.1. Chapter Overview

This chapter provides a global view of the architecture proposed to address the challenge of dynamic federation. The description starts in Section 4.2 with a brief introduction to system modeling definitions and basic concepts that contextualize the methodology.

After this introduction, a general architectural model that is common to FIM systems is presented as a reference to introduce our extensions. Based on this basic model, a requirement analysis is performed.

Next, Section 4.3 introduces the contributions made in this thesis to extend the functionality of the basic architecture satisfying the stated requirements. The new components and the extended functionalities are explained and accompanied by a flowchart that illustrates how each architectural component interacts with the rest.

Finally, Section 4.4 ends with the main conclusions extracted from the architecture definition procedure.

4.2. Design Preliminaries

As introduced in the previous chapters, the focus of this thesis is on contributing to make federated identity management systems more dynamic. For this purpose, an essential first step is developing a system model. Firstly, in this preliminary section, we set the basis for the model description following two steps: 1) introducing the concepts and definitions required to contextualize the modeling methodology; and 2) describing the general architectural model supported by current identity management systems as a basis to derive the requirements for a dynamic model that will drive our proposal.

4.2.1. System Modeling Definitions and Concepts

A system model constitutes the conceptual model that formally describes and represents a system. When describing a system model in the IT field, the software architecture discipline is commonly used. The software architecture discipline is centered on the idea of reducing complexity through abstraction and separation of concerns. Software architecture as a concept has its origins in the research of Edsger Dijkstra in 1968 and David Parnas in the early 1970s, who emphasized that the structure of a software system matters and getting the structure right is critical. The study of the field increased in popularity since the early 1990s with research work concentrating on architectural styles (patterns), architecture description languages, architecture documentation, and formal methods [Garlan and Shaw, 1994]. However, to date there is still no universal agreement on the precise definition of the term "software architecture", as pointed out in the study performed by the Carnegie Mellon University [Carnegie Mellon Software Engineering Institute (SEI), 2006]. Some classical definitions of the term gathered in this study are:

- *"The structure of the components of a program/system, their interrelationships, and principles and guidelines governing their design and evolution over time."*- Garlan and Perry, IEEE Transactions on Software Engineering, 1995.

- *"A software system architecture comprises: A collection of software and system components, connections, and constraints; A collection of system stakeholders' need statements; A rationale which demonstrates that the components, connections, and constraints define a system that, if implemented, would satisfy the collection of system stakeholders' need statements."* - Boehm et al. USC Center for Software Engineering, 1995.

- *"An architecture is the set of significant decisions about the organization of a software system, the selection of the structural elements and their interfaces by which the system is composed, together with their behavior as specified in the collaborations among those elements, the composition of these structural and behavioral elements into progressively larger subsystems, and the architectural style that guides this organization—these elements and their interfaces, their collaborations, and their composition"* - Kruchten, The Rational Unified Process. Also cited in Booch, Rumbaugh, and Jacobson, The Unified Modeling Language User Guide, Addison-Wesley, 1999.

More recently, modern definitions of the term appear in *"Documenting Software Architectures: Views and Beyond* [Clements et al., 2003]", *"Software Architecture in Practice"* [Bass et al., 2003], and *"ANSI/IEEE Std 1471-2000, Recommended Practice for Architectural Description of Software-Intensive Systems"* [Jen and Lee, 2000]. All of them are closely related and overlapping, and thus contain common points that can be summarized. A simple and understandable definition, mainly based on [Bass et al., 2003], is:

"An architecture description is a formal description and representation of the set of structures needed to reason about the system, which comprises components, relationships between them, and properties of both."

The primary goal of architectural modeling is coming to a representation or understanding with respect to how to build a system. The presented definitions put emphasis on the fact that systems are comprised by several components and there are relationships, dependencies and connections between them. Due to this complexity, a common issue in architectural modeling is the need for various architectural views. A view is captured as a combination of diagrams and text descriptions (such as use cases, technical specifications, or prose). Typical views include, e.g., logical, usage/business process, user interface, deployment, data storage/transmission, etc., though there is not a consensus in regard to the set of views that is required for a project. Not all views are relevant to all systems. Instead, they should be chosen depending on the nature and context of the project. In this sense software architecture is really the amalgamation of the multiple perspectives a system always embodies.

There are also formal languages in order to describe all the views that comprise a software architecture. These languages are called ADLs (Architecture Description Languages), and several of them have been devised. Despite no consensus exists on which symbol-set or language to be used for each architectural viewpoint, the UML (Unified Modeling Language) is a commonly used standard [Fowler, 2004]. It is defined as a language that can be used *"for analysis, design, and implementation of software-based systems as well as for modeling business and similar processes."*

Having introduced the above definitions and concepts, we have the basis for describing our architectural proposal. In presenting the architecture, we follow the definition in [Bass et al., 2003], and we use UML diagrams and flowcharts to depict the views. We use a twofold view approach encompassing:

- A **Logical view**, which describes how the system is structured in terms of components or units of implementation. It shows dependencies between elements, interface realizations, part-whole relationships, and so forth.

- A **Process view**, which explains the system processes and how they communicate.

With the first view we cover the static or structural part of the system, while the second view depicts the dynamic or behavioral part. Diagrams will be always accompanied by descriptive text in order to complete a whole description of the system.

4.2.2. Requirements Analysis

Now that we have defined the methodology to be followed in the description of our model, we can continue with the definition of the architecture. For this purpose, we first analyze the current architecture of FIM systems, since it will be the starting point to add the new functionality.

In this regard, there is not a single FIM architecture definition. The only thing that can be found in the literature are standards, specifications, as well as guidelines and best practices for implementation. But every organization deploying a FIM solution defines its own architecture. In the specific case of SAML, which is the most important and complete framework, there are two big implementations: Shibboleth in education environments and SAML-based deployments in industrial/governmental environments. By extracting the common features of them, we have elaborated an UML component diagram for a generic identity federation architecture, depicted in Figure 4.1.

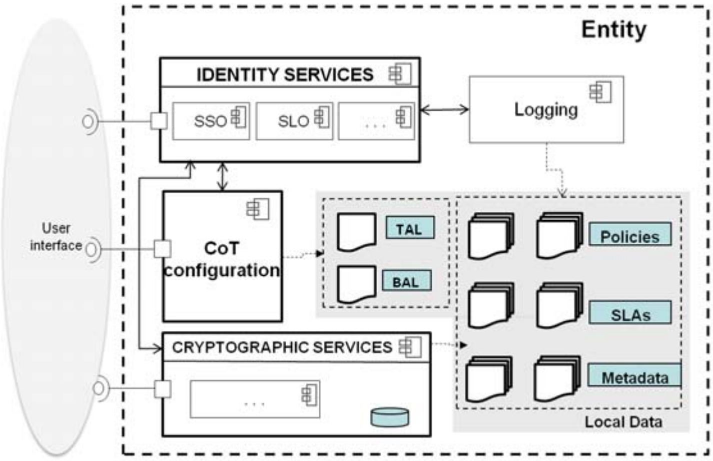


Figure 4.1: Generic architecture for identity federation

We use generic names for the components in Figure 4.1. However, it is worth noting that modules may be called differently across implementations despite their functionality is the same. As it can be observed in the image, the architectural components in the providers are:

- **CoT Configuration component**. Both SPs and IdPs implement a Configuration component, over which the services rely. This component is in charge of accessing local data stores to determine if a provider involved in a current identity-related transaction is trusted. This decision is made basically by consulting the local data stores to check if the entity is contained in a list of trusted entities and, if necessary, if an explicit SLA exists between them. An external interface is offered to administrators in order to configure the static trust lists previous to interaction, a procedure that is performed manually. Automatic inclusion of providers in the trust list is possible in some cases, but in a limited way. More details will be given about these procedures in Chapter 6.

- **Identity Services component.** This component encompasses the services offered by the identity framework, i.e., Single Sign-on, Single Log Out, Authentication, Authorization, etc. Both kind of providers implement this component to allow the communication and exchange of user identity data between them.
- **Cryptographic Services component.** This component gives the cryptographic support for the security processing required, i.e., encryption/decryption of assertions, signing/validation, etc.
- **Logging component.** Providers usually implement a Logging module in order to monitor user and service activities. The registries are used by the identity services, but an interface may be also provided for auditors (external parties).
- **Data store**. It contains information used by the rest of the components, i.e., metadata documents, policies, SLAs, trust data, credentials, logs, session data, messages, etc.

As it can be noted, the service and trust layers of the FIM model presented in the introductory chapter in Figure1.1 are implemented by the above components[1] . However, the trust layer functionality is limited, because it is just implemented as a configuration module where relationships are inserted statically. Since the trust layer is the basis to achieve flexible models and allow scalability, our main goal in this thesis is to extend and enhance this configuration component and make it evolve towards a more complete component. It is important to note that we contribute on a specific part of the architecture, but without losing the global perspective.

[1] for the sake of simplicity, we do not include the architectural components that cover the location and translation layers. These topics are out of the scope of the research presented in this thesis, and so the existence of components that perform location and translation functions is assumed

Thus, having the generic architecture as a reference starting point, the next step is the derivation of requirements to design an architecture that allows dynamic federation, i.e., automatic creation and maintenance of secure trust relationships. For this purpose, as developed during the past chapters, we will introduce reputation data and risk assessment. Reputation can be used as a basis for trust when the potential collaborator is unknown, then the evolution of the trust relationship will be based on the direct interaction or transaction history. In turn, risk evaluation gives an idea of the probability that the collaboration results in an undesired output. In order to identify the requirements for designing a suitable reputation/risk model to integrate in these scenarios, we first divide FIM into two different phases, as depicted in Figure 4.2:

1. **Pre-Federation Phase (or Federation between providers).** This phase consists of the establishment of a relationship between providers: deciding on protocols to interoperate, agreeing on common rules and policies, etc. It can be understood as a Bootstrapping Phase, which allows parties to gather information about each other and to initiate cooperation. The SLAs that govern the relationship are negotiated in this phase.

2. **Post-Federation Phase (or Transactions between federated providers).** This phase encompasses transactions between two federated entities (e.g., requesting user attributes or accepting authentication claims). At this point, entities have basic information to support their decisions. At least, the deciding entity will have data derived from the Bootstrapping Phase and, if more interaction has taken place, it will also maintain a history of transactions. It can be viewed as the Evolution Phase, since entities progressively construct and consolidate their relationships.

In current FIM systems, Pre-Federation consists of establishing agreements between entities in the federation and manually setting up a closed Circle of Trust. Then, in Post-Federation it is assumed that everything will work, since interaction is only possible with the entities that were pre-configured in the previous phase. Our goal is to allow entities to dynamically move from Pre-Federation to Post-Federation with a certain degree of trust and continually monitor and consolidate this trust relationship. We base our proposal on the inclusion of external information (reputation) and the computation of trust and risk values upon which decisions to cooperate can be made. For this purpose, the main requirements we identify to operate in these two phases are:

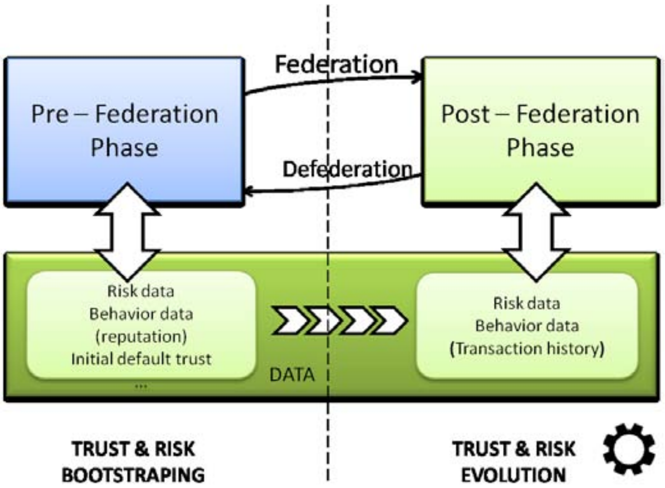


Figure 4.2: Pre-Federation and Post-Federation phases in FIM

- Local computation of trust values
- Local computation of risk values
- Storage, dissemination and aggregation of reputation data to be included in the trust computation procedure
- Dynamic decision making based on trust and risk

In addition, in order to maintain and monitor the evolution of relationships in Post-Federation, another functionality is also required:

- Monitoring and adjustment of trust levels: transaction history, trust update, SLA conformance evaluation, etc.

Basically, the goal in Pre-Federation is to make a decision whether to federate or not with the other entity and to which extent. To make this decision, information must be gathered regarding technical and operational issues, but also regarding reputation. In the case of Post-Federation, the requirements are the same as in the previous phase but oriented to make decisions about particular transactions instead of deciding whether to federate or not. For this reason, additional mechanisms are required to monitor the current state of the trust relationship and decide whether to give more privileges, deny some of them or terminate the federation.

In order to cover the above requirements we present a solution for federation based on trust and risk. The detailed explanation of the proposed architecture is developed in the next section.

4.3. Architecture Description

4.3.1. Components and Relationships

According to the requirements presented above, Figure 4.3 shows the architectural components to be included in those entities that take part in the dynamic federation model, namely IdPs and SPs.

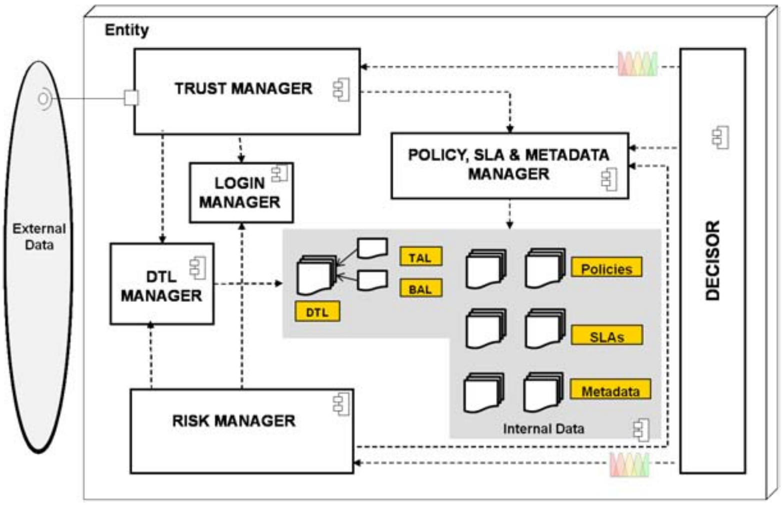


Figure 4.3: Proposed architecture for dynamic federation between SPs and IdPs

Next, we explain the details of every element in the presented architecture, and how they extend and enhance the functionality provided by the components in the generic architecture. We consider two kind of *Data* (external and internal) and a set of components denoted as *Managers*, which implement the logic to handle those data and allow dynamic federation. Thus, the components are:

1. ***Internal Local Data***

   Each entity handles the following internal data:

   - **Dynamic Trust List (DTL).** In current implementations of FIM frameworks it is usual that entities are configured with static lists of trusted entities. More specifically, SAML-based deployments implement Trust Anchor Lists (TALs) and, in some cases, also Business Anchor Lists (BALs), that must be preconfigured before any interaction between parties takes place. These lists contain the digital certificates associated to every other entity that is considered trustworthy. Protocol messages whose digital signature cannot be validated against the TAL/BAL are rejected. Thus, trust does not evolve over time, because interaction experience is not taken into account, community knowledge is not exploited, distrust and ignorance are treated in the same way, and the automatic establishment of trust relationships between unknown entities is impossible. The pre-configured TAL/BAL model poses important obvious limitations in dynamic open environments. So, instead of static lists, the system maintains an enhanced Dynamic Trust List with more complete information. The DTL stores trust, reputation and risk data regarding other entities in the FIM infrastructure and it is automatically updated according to the establishment and evolution of trust relationships (see Chapter 6 for more details on the DTL content). Furthermore, in order to maintain compatibility with existing deployments and to allow the establishment of relationships based on previous agreements, TALs and BALs can also exist and, when this is the case, they will be used to initialize the contents of the DTL. Also, an external interface for administrators to consult and/or modify the DTL content is provided through the DTL Manager, explained below.

- **Policies, SLAs and Metadata**. Entities must define policies regarding different aspects of FIM, such as the supported cryptographic algorithms, thresholds and rules for risk, trust and reputation values, etc. Besides the policies, each entity defines SLAs to describe the extent of federation relationships, e.g., the service availability, the minimum level of assurance to be used in authentication, etc. Finally, each entity will have metadata documents, which are used to specify the technical information required to configure a federation relationship using a particular federation framework (e.g., SAML Metadata [Cantor et al., 2005c]): supported federation protocols and use cases, digital certificates to be used in communication, etc.

In regard with the internal data, the novelty lies on the extension of the information stored, which will be used as source to make decisions. With the new DTL data structure and by extending the data in the local policies, SLAs and metadata to include trust, reputation and risk related info, the storage requirements of our model are satisfied.

2. ***External Data***

   In addition to the internal local data, entities need to obtain external information in order to enrich their knowledge about the other entity in a transaction. Thus, the external information to be gathered is:

   - **Policies, SLAs and Metadata**: When deciding about to federate or not with another entity it is required to obtain these data to compare with the local information and ensure that federation is technically, operationally and legally feasible to some desired extent. If federation is not possible in this sense, then there is no need to ask for reputation data to make a decision.

   - **Reputation Data**: If the entity is unknown, reputation data will be requested in order to compute an initial trust level to make the decision whether to transact or not. The mechanisms to obtain and include reputation data in the trust computation procedure will be explained in Chapter 6.

   These external data, once obtained, are stored locally and handled by the related managers. The storage of data regarding cooperating providers is also contemplated in the original generic architecture. The novelty here is the possibility to gather reputation data as a source to calculate the trust to be placed in a collaborating entity on the fly.

3. ***Managers***

   In order to handle internal and external data (CRUD operations: create, read, update and delete) and implement the logic for dynamic federation, the following components are introduced:

   - **DTL Manager.** This component is in charge of handling DTL operations: create, read, update and delete data. Essentially, it communicates with the *Risk* and *Trust Managers* and performs the operations required by them. The values contained in the DTL are used to make the trust and risk computations.

The list is dynamically updated under specific events, e.g., when receiving recommendations from other entities or when a successful interaction ends.

- **Policy, SLA and Metadata Manager**. This component is used by the *Trust*, *Risk* and *Decision Managers.* It is in charge of the following main functions:

  a) Communicating trust related rules to the *Trust Manager* module (e.g., thresholds for malicious entities, default trust values, etc.).

  b) Communicating risk related rules to the *Risk Manager* module (e.g., risk thresholds, minimum requirements, etc.).

  c) Performing CRUD operations over the local data, e.g., if a trust relationship evolves positively, the *Trust Manager* will notify this module in order to extend the initial SLA and grant more permissions to the transacting entity.

  d) Providing policy information to the *Decision Manager*, so it can combine the computed trust and risk values and make a decision.

- **Logging Manager**. This component is in charge of registering historical information for auditing purposes. Apart from this functionality, it is required that it monitors and registers the bad and good actions of the providers in order to have an history of transactions for reasoning and deriving risk and trust values.

- **Trust Manager**. This component is in charge of processing external and internal trust information to compute a trust value related to a collaborating entity. It reads data from the *DTL Manager*, *the Policy, SLA and Metadata Manager*, and provides an interface to obtain data from external sources. Its functionalities include: execution of a reputation (query/response) protocol to obtain data about unknown entities, aggregation of opinions and calculation of trust values.

- **Risk Manager.** This component evaluates the risk attached to the current transaction. It receives data from the *DTL Manager*, *the Policy, SLA and Metadata Manager.* From these sources, the component obtains all the required input information to compute risk.

- **Decision Manager**. This component is in charge of deciding whether to initiate or not a transaction with another entity, in case of being a requestor; or whether to respond or not to a transaction request. The inputs for this module are the computed trust value, the risk associated to the transaction and the internal policies that will be used to govern the decisions. All these data are obtained through the related managers.

The traditional architecture also provided means for accessing policies, metadata, SLAs and trust lists through a Configuration component and by means of an administrative interface. Here we have split this functionality into different dedicated components. But the novelty lies on two main pillars: 1) the extension of the trust functionalities in order to allow the inclusion of reputation data and the computation of trust based on it; and 2) the introduction of a completely new module for risk assessment. Furthermore, thought not shown in the picture in Figure 4.3, interfaces are also provided for administrators to consult/modify data. Also, the Identity Services and the Cryptographic Services components are not shown because we do not add anything new to them, but they are also part of the architecture.

The *Trust* and *Risk Managers* may be fully implemented or not. For example, an implementation may include only a *Trust Manager* with the reputation functionality, or only a *Risk Manager* combined with the original static trust management functionality. The combination of risk and trust values is performed by the *Decision Manager*, which is also a new component.

Finally, the *Logging Manager* simply extends the original Login component to register information about the satisfaction of transactions, in order to aid in the trust and risk computation.

With these set of managers, the design requirements listed in the previous section are fulfilled.

Summarizing, since the trust and risk components include the functionality that constitutes the main contribution of the thesis, a separate chapter is dedicated to define each of these components. Thus, Chapter 5 and Chapter 6, develop the risk and trust models to be implemented as part of the architecture. Then, Chapter 7 shows how the decision manager combines the outputs of the trust and risk managers for decision making.

Finally, in order to complete the general picture of the architecture, next section explains the general behavior of the architectural components through an operation flowchart.

4.3.2. Operation Flow

In order to show the behavioral part of the architecture, this section illustrates how the different components interact with each other.

In a dynamic federation environment there can be two interaction scenarios for a provider:

a) *Transacting with a federated provider.* In this case both the SP and the IdP belong to the same federation and they know each other from previous interactions, so trust information is available in their DTLs.

b) *Establishing a federation with an unknown provider.* In this case, since entities are unknown to each other because no previous interaction exists, there is no information in the DTLs.

Next we explain the operation flow (Figure 4.4) that covers the above cases, referencing which parts of the proposed architecture take part in the process. The aim is to provide just a conceptual understanding so low level details are not yet given, but they will be addressed in subsequent chapters.

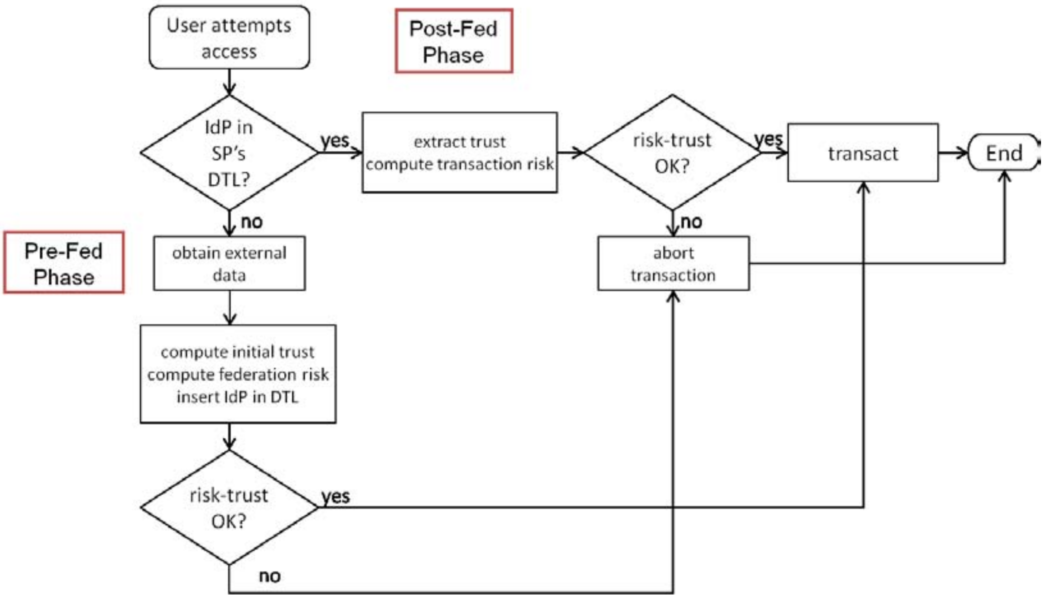


Figure 4.4: Flowchart for interaction using the dynamic federation architecture

On the one hand, the steps followed in the case of known providers are:

1. First, a user attempts to access a service offered by a SP.
2. The SP discovers the IdP that the user wants to use for identification and checks its DTL to determine if the IdP is contained in it.
3. If the IdP is contained in the DTL, this means that entities are already federated, configured to operate and a trust value exists. So the next step consists on extracting the trust value, computing the risk value associated to the transaction and combining both.
4. If the combined trust-risk value is enough for cooperation, then the SP will continue with the transaction. On the contrary, the provider will not continue with the transaction.

On the other hand, when the providers are not federated, just the first two steps are the same as in the previous case. But when checking the DTL:

3. If the IdP is not contained in the DTL, this means that entities are not federated and the relationship must be established. Thus, the next step consists on obtaining external data (i.e., reputation, metadata, etc.) about the IdP.
4. Next, the gathered external data are used to compute trust and risk values, and the IdP is inserted in the DTL of the SP together with this information.
5. If the combined trust-risk value is enough for establishing a federation, then the SP continues with the transaction. On the contrary, no federation is established.

In the presented cases, we show the operation from the point of view of an SP that wants to transact with an IdP. Although not reflected, the IdP on the other side also checks its DTL and performs the same steps to determine if it should transact with the SP.

As far as the architectural components involved at each point of the operation flow, Step 2 of both scenarios is performed by the *Trust Manager* through the *DTL manager*. In both cases, the calculations to obtain the trust, risk and combined values, are performed by the *Trust*, *Risk* and *Decision Manager*, respectively. This happens at Step 3 in the federated case, and Step 4 in the unknown scenario. The gathering of external data performed in Step 3 of the unknown scenario is implemented by the *Trust Manager*, which uses the DTL and the *Policy, SLA and Metadata Managers* to store the information. Finally, the comparisons with the decision thresholds for the combined trust-risk value in the last steps of both cases are performed by the decision manager.

Finally, after presenting this general view of the architecture, we will develop the functionality and low level details of the main modules in the next chapters.

4.4. Conclusions

It is clear that current FIM architectures are limited to provided secure and dynamic means to establish relationships between providers. In this chapter we proposed an extended architecture to fill this gap. The architecture is composed of a set of logical modules that separate and encapsulate the functionalities required to achieve dynamic federation. The pillars of the architecture are the risk and trust components, which constitute the main contribution of the thesis. The mathematical models implemented by each part are thus developed in the following chapters.

In conclusion, the extension of the architecture satisfies the intended goals, since it makes possible to minimize pre-configuration requirements by allowing relationships to be established on demand based on trust and risk analysis.

Chapter 5

Risk Model Proposal

"If you cannot measure (or model) it, you cannot improve it"

Lord Kelvin

Contents

5.1. Chapter Overview

This chapter describes the proposed risk model that is to be included by participants in federated identity management scenarios. The general goal of the model is to serve as a tool for decision making that:

- provides a meaningful numerical value that condensates risk information
- allows entities to include subjective preferences according to their interests
- aids entities in deciding whether to collaborate or not with another entity
- aids entities in deciding which entity in a set is the best alternative for cooperation

Between the different types of risk assessment methods (see Chapter 2), we choose to follow a semi-quantitative approach. The reason is that pure quantitative analysis is not possible since no statistical data are available to build the model. Qualitative scales are thus required, but instead of staying purely qualitative, the mapping of these scales to quantitative values permits the automation needed for dynamic federation. Thus, the hybrid approach is the one that best fits our needs.

The methodology employed to define the risk model consists of three steps. Firstly, we design a taxonomy to capture the different aspects of a relationship in identity federation that may contribute to risk. This approach allows us to decompose the complex problem of risk assessment and to acquire a detailed knowledge. Secondly, based on the taxonomy and aiming at developing a computational model, we propose a set of metrics as a basis to quantify risk. Finally, we describe to aggregate the metrics into a meaningful risk figure, coming to the final formal definition of the model.

According to the stated goals and methodology, the chapter is divided in three sections which develop the taxonomy, the metrics and the aggregation procedure, respectively.

5.2. Towards a FIM Taxonomy for Risk Metrics Derivation

5.2.1. Why a Taxonomy?

"Divide et Impera"

The exact definition of taxonomy varies slightly from source to source, but the core idea behind the discipline remains the same: identification, naming, and classification. The history of taxonomic classification is deeply rooted on biology, where theory and practice of grouping individuals into species has been crucial for the understanding of biodiversity and conservation. But the use of taxonomies has proved useful in many other fields, such as education [Anderson et al., 2000] or psychology [McGarty, 1999]. Indeed, anything may be categorized according to some taxonomic scheme: animate objects, inanimate objects, places, concepts, events, properties, and relationships.

As a general point of reference, the Oxford dictionary [1] defines taxonomy as a "*scheme of classification*". Furthermore, Flood and Carson [Flood and Carson, 1993] point out that a taxonomy serves several purposes:

- **Description:** It helps us to describe the world around us, and provides us with a tool with which to order the complex phenomena that surround us into more manageable units.
- **Prediction:** By classifying a number of objects according to our taxonomy and then observing the 'holes' where objects may be missing, we can exploit the predictive qualities of a good taxonomy. In the ideal case, the classification points us in the right direction when undertaking further studies.
- **Explanation:** A good taxonomy will provide us with clues about how to explain observed phenomena.

In the light of the utility of taxonomies to approach complex problems, we aim to make a classification of the different aspects involved in establishing and cooperating in a federation between identity and service providers. The main goal, oriented to contribute towards risk modeling, is to use the taxonomy as a tool that helps in the derivation of metrics for risk assessment. Understanding risk as a complex multidimensional construct, as we already clarified in Chapter 3, the categorization of different contributor factors is crucial as a first step to find the pieces that will conform the overall risk.

Taxonomies can be designed in different ways, being hierarchies the most common structures. In a hierarchical taxonomy, categories or nodes progress from general to specific. Thus, each subsequent node is a subset of the higher level node. We followed a hierarchical approach because, once the metrics for each category are derived, risk can be easily calculated following a hierarchical aggregation model directly grounded in the taxonomy.

Summarizing, we aim to design a generic taxonomy that gathers the features of contemporary federations, but also that allows to identify 'holes', i.e., where to place adequate

[1] http://oxforddictionaries.com/

metrics to be used for risk calculation in the dynamic federation scenario envisioned in this thesis. For this purpose, we have studied all the requirements taken into account when establishing a federation according to the different identity management specifications (see section 2.1) and organized them in categories. The whole rationale design and justification of each category in the taxonomy is explained through the following section.

5.2.2. Rationale Design

This section goes through the proposed taxonomy justifying how we made every design decision to model and classify the different categories in FIM where risk metrics may be grouped.

First of all, based on the understanding of FIM as a two-phased procedure (idea proposed in Chapter 4), we first divide the taxonomy in two main categories: **Pre-Federation** and **Post-Federation**. Since the justification of this division has been already well-argued, no further elaboration is provided here. It should be clear that the decisions to make and the available information are different in each phase; and so are the faced risks:

- In **Pre-Federation**, an entity will presumably have to decide whether to establish a relationship for further cooperation with another entity, i.e., to federate or not to federate. The sources of information to compute a risk metric at this stage could be, for example, the entity metadata, pre-configured relationships with other entities, internal policies, and the Service Level Agreements being negotiated.

- On the other hand, entities in **Post-Federation** phase will have to decide whether to transact or not in the context of a particular service. Metrics to assess risk at this stage can be calculated from the information available in the assertions and protocols in use, the characteristics of the specific service transaction in process, and, also, by leveraging the risk metrics from the previous phase.

The distinction of these two fundamental phases in FIM constitutes the first level (L1) of the proposed taxonomy, whose complete schematic is depicted in Figure 5.1. In the following subsections, we explain and argument the different subcategories related to each phase.

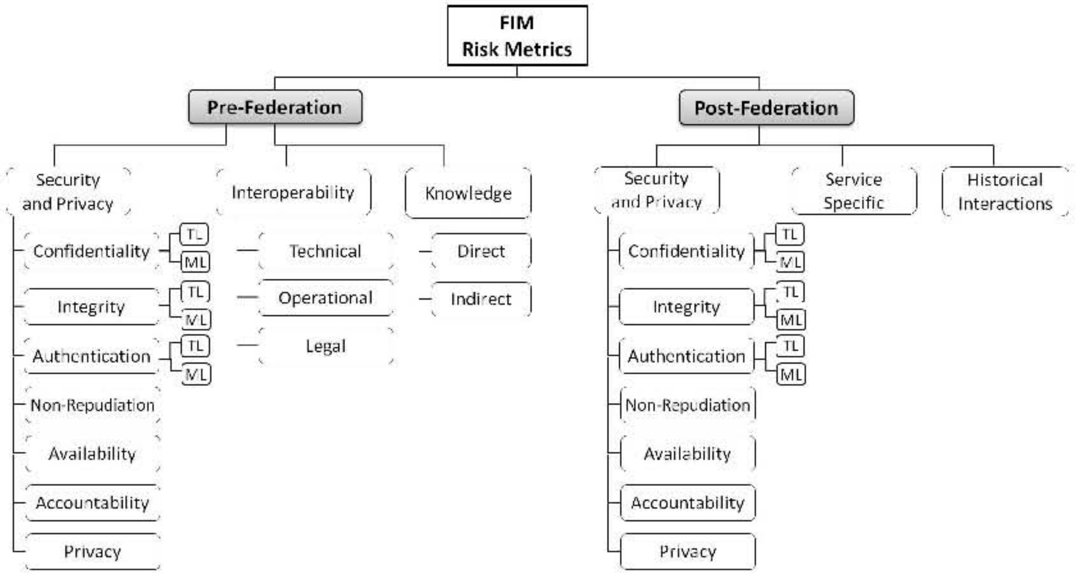


Figure 5.1: Taxonomy for risk metric derivation in FIM

Pre-Federation Risks

Our approach to categorize risk metrics in Pre-Federation stage starts with finding the answers to these questions:

1. *What is required to move from Pre-Federation to Post-Federation?* or, equivalently, *What are the requirements to establish a federation?*

2. *Which are the different aspects of these federation requirements that may contribute to the global risk?*

Consequently, in order to determine the different subsets to categorize Pre-Federation risk metrics, we reviewed the different documents used to establish contemporary federations.

After the analysis, we saw that to establish a federation between providers, a set of agreements must be put in place. For this purpose, there is not a standard or common fixed set of minimum requirements; instead every deployment follows its own rules based on the federation framework in use, and on the goals and purpose of the federation itself. Some identity management frameworks, such as SAML and Liberty ID-*, provide additional documents apart from the core specifications that describe a set of guidelines for federation and also recommend best practices. However, other identity frameworks are not specific in this regard and simply assume in their general specifications that cooperating entities are configured to be able to interoperate and trust each other.

Since the documents for federation establishment provide guidelines that are subject to different interpretations, we aimed also (for the sake of completeness) to analyze how federations are implemented in practice, to which extent are the agreements put in place and which features are covered in these agreements. Nevertheless, there is no much public information about how federations agreements are deployed in the real world, except for the survey of Research and Education Federations (REFEDs) in [TERENA, 2012], carried out by the Trans-European Research and Education Networking Association (TERENA).

In examining both the existing specifications and the real world public data about federation establishment some common trends are worthy of mention:

1. In all the identity frameworks, **security and privacy** considerations are taken into account to some extent.

   SAML core documents recommend to apply digital signature mechanisms to protocol messages, and assertion encryption is supported as well. Furthermore, there is a complete document dedicated to cover SAML security and privacy considerations [Maler et al., 2005]. This document highlights the importance of applying mechanisms to achieve confidentiality, integrity and authentication at the transit and message level. For this purpose, it recommends also to use SSL 3.0 [Freier et al., 2011] or TLS 1.0 [Dierks and Rescorla, 2006] to secure the communication channel. Likewise, the Liberty "*Deployment Guidelines for Policy Decision Makers*" [Varney and Sheckler, 2005] addresses certain privacy and security related considerations that federation participants should consider. Similarly, WS-Federation specifications require agreement between parties on security claims and also state that agreement on mechanisms for securely transporting those claims over unprotected networks may be required. OpenID, OAuth and Information Cards also include security notes in their specification documents and emphasize, for example, that the use of encryption and digital signature mechanisms has an impact in the probability of being attacked.

   On the other hand, security considerations are also present in most of the public federation agreements in the REFEDs survey, specially regarding to privacy of user personal data.

   Consequently, security must be definitely taken into account when assessing federation risks, since inappropriate protection may lead to attacks. The understanding and agreement on a common set of security practices, which ensure that every entity works within a desired risk context is necessary.

2. There is a clear need for **interoperability**.

   At least entities must be able to communicate, i.e., to use the same protocols. Liberty ID-* and SAML require that metadata is exchanged in order convey configuration information required to interoperate. WS-Federation also requires the exchange of metadata as a basis for interoperability. Furthermore, based on WS-Policy, web services are allowed to use XML to advertise their policies on different issues (e.g., quality of service) and for web service consumers to specify their policy requirements, guaranteeing thus interoperability. Apart from the specifications, Liberty also provides a guidelines document [Sheckler, 2007] which recommends several relevant aspects that should be addressed to ensure robust relationships between participants in a federation. More specifically, the document points out the importance of agreements on operational rules, technical standards for communication and applicable laws according to the jurisdiction(s) in which the federation operates and the class of business. In the case of OpenID, OAuth or Information Cards technologies, their core specifications assume that interacting parties are technically interoperable (i.e. configured to understand the same set of protocols). Finally, the agreement upon different issues related to interoperability is also reflected in public policies that govern the federations surveyed in [TERENA, 2012].

   We can thus conclude that interoperability has an impact in the risk when connecting entities in a federation.

3. The required degree of **knowledge** regarding the other party involved in a federation transaction may vary.

   Depending on the identity framework, knowledge about the transacting entity is required to interact or not. In those cases where knowledge is required, it may be pre-configured (direct) or transitively obtained (indirect). And this knowledge is used to place trust as it was explained in Chapter 3. In this regard, SAML,

> Liberty and WS-Federation provide well-defined models to derive knowledge based on pre-configuration and using transitive relationships. In "*OASIS Trust Model Guidelines*" [Boeyen et al., 2004], a classification of possible implementation models is presented based on direct/indirect business knowledge and on direct/indirect authentication knowledge. On the other hand, OAuth and Information Cards simply state in the specifications that the interacting entities should be known to each other. Typically knowledge is maintained on lists of known entities, which are manually updated by administrators. In turn, the OpenID framework considers that interaction can happen even without the need for previous knowledge following a *"trust-and-accept-all-comers"* philosophy. And finally, the federations surveyed by TERENA require new entities to sign a membership agreement previous to enter the federation, so the knowledge is based on direct contractual frameworks. We conclude that knowledge influences risk, since it is vital to determine a trust level.

Based on this thorough analysis of FIM specifications, related best practices and recommendations, and the reviewing of public data about how current federation deployments were established, we conclude that the whole set of aspects to be considered before creating a federation are oriented to achieve security and privacy objectives, to establish interoperability rules and policies for legal, technical and operational compliance, as well as to determine the knowledge regarding the interacting party. The taxonomy proposed here aims to cover all these aspects in order to be generic enough to abstract all the federation frameworks. Accordingly, we divide the next level (L2) of classification in the Pre-Federation phase into three main blocks or categories, namely: **Security and Privacy**, **Interoperability** and **Knowledge**.

Next, each of these L2 classes is explained in more detail and further divided into subcategories.

The first class, Security and Privacy , encompasses those risk metrics related to the security and privacy features that are supported by an entity who wants to establish a federation. These features, which constitute the subcategories located at level L3 in the taxonomy, are:

- **Confidentiality**: disclosing information only to intended and authorized recipients.

- **Integrity**: guarding against improper information modification or destruction.
- **Authentication**: confirming something (or someone) as authentic, verifying the validity of the claims made by or about the so called subject.
- **Non-Repudiation**: provides evidence that one party involved in a transaction sent or received a message, so it cannot be denied.
- **Availability**: ensuring that a system is operational and that it is accessible to those who need to use it, so the business purposes can be met; loss of availability is often referred to as “denial-of-service”.
- **Accountability**: the ability to associate a consequence with a past action of an individual. It is required that the individual can be linked to action or event for which he/she is to be held accountable.
- **Privacy**: appropriate use and protection of information, which means seclusion and selective disclosure of data according to law and policies. Privacy is sometimes related to anonymity, defined as the wish to remain unnoticed or unidentified in the public domain.

The above categories were chosen based on the features commonly considered in the security literature.

Each of the above basic security services or “CIA” (i.e., Confidentiality, Integrity and Authentication) depends on the cryptographic characteristics of both the data exchanged at the message level (ML) and at the transport layer level (TL). As mentioned before, most of the FIM protocols strongly recommend the use of secure communication protocols, such as Secure Sockets Layer (SSL). Thus, it is required to evaluate the quality of the security services that can be provided at the message level and also with regard to the communication nature. For example, a FIM transaction with encrypted assertions that are also transmitted over a secure SSL connection would incur in less risks than if the communication channel is not secured. This requirement is also reflected in the taxonomy, as a sub-classification (at level L4) to be evaluated for each basic security service.

The second subclass under Pre-Federation, called Interoperability, encompasses those issues related to interworking between entities. Interoperability can be decomposed in three different domains, located at level L3 in the taxonomy: **Technical**, **Operational** and

Legal. The Technical category is required since there are many different technologies for FIM. Furthermore, inside a specific framework there can be different implementations. For example, two SAML enabled entities could not interoperate if they do not support a set of common Profiles. Thus, metrics are required in order to measure the technical compatibility of the systems. Apart from technology related interoperability, it is also important to evaluate if the policies of each entity are compatible to a certain degree. In these sense, entities should measure, on the one hand, the interoperability regarding Operational policies. On the other hand, legal compliance and applicable jurisdictions should be also addressed, so cooperation is also interoperable at regulatory level. Therefore, if cooperation implies risk of violating an entity policy, it should be avoided. Thus, the interoperability metrics are computed based on the information gathered from metadata, SLAs and policies of the other entity before deciding on interaction.

Finally, the last differentiation to allocate risk contributing factors in Pre-Federation is named Knowledge and involves those risk factors related to the previous information that is known about the other entity. Risk metrics derived from this category will thus aid in the quantification of the initial trust level that is assigned to the interacting party, being a point of relation between trust and risk. Since each party has to decide whether to engage in a relation with other party for future cooperation, it is reasonable to gather information about its trustworthiness. In order to capture the different type of relationships, Knowledge metrics are sub-classified into **Direct** and **Indirect**. The Direct Knowledge metrics are related to pre-configured relationships (e.g., digital certificates or business agreements); whereas Indirect Knowledge refers to the reputation data, information that can be obtained from external sources [Gómez Mármol et al., 2010]. Consequently, the risk level will vary according to the existing knowledge, because it is indirectly proportional to the uncertainty.

One could argue that the taxonomy could be organized in another way, and it could be indeed. For example, security could be included as a category under interoperability, since it can be understood that entities are not interoperable regarding security objectives if their policies on algorithms are not compatible. However, the design approach does not limit the utility of the taxonomy as a tool for modeling aspects of federation relationships and derive metrics for risk assessment. Furthermore, despite no other taxonomy has been proposed to compare with our model, our classification is coherent with the roadmap for the study of trust federations proposed by Kantara[2].

Post-Federation Risks

Here we follow the same methodology used for identifying the possible risk contributions in Pre-Federation. Thus, we try to answer these two questions:

1. *What kind of transactions are performed in Post-Federation?*
2. *Which are the different aspects of Post-Federation transactions that may contribute to the global risk?*

According to the specifications of the different identity frameworks, after a federation between providers has been established, transactions to exchange user identity data can be performed between them. In this phase, risk must be assessed on a per-transaction basis. Some facts are relevant at this phase: 1) firstly, the requirements and agreements for interoperation have been already put in place so interoperability is assumed under a well-defined Federation SLA; 2) depending on the context of the transaction (i.e., characteristics of the service, required personal information, etc.), security and other service specific features must be assured to different extents, 3) transacting entities will have additional knowledge based on previous transactions apart from the initial pre-configured knowledge data. Based on these facts, as shown in Figure 5.1, Post-Federation Risk metrics are categorized in three main classes (level L2), namely: **Security and Privacy**, **Service Specific** and **Historical Interaction**.

Next, each of the L2 classes under Post-Federation is explained in more detail.

In Post-Federation transactions it is relevant to measure how the security features are fulfilled in order to know how they contribute to the total risk. Thus, the same classification of Security and Privacy made for Pre-Federation applies here. The difference is that the Security and Privacy metrics taken in Post-Federation are related to the current transaction and so they are used to decide whether to transact or not in the measured conditions. However, metrics in the Pre-Federation phase are used to determine the global support of security features and decide whether to federate or not.

[2] http://kantarainitiative.org/confluence/display/TFMMWG/TFM+Topic+Map

Apart from the Security and Privacy related metrics, we consider a further distinction in order to include Service Specific Risks and ensure completeness. This category allows the risk model to be tailored for different types of services. Services differ in characteristics such as required personal information, value of the resources owned by the SP, importance of data availability and so forth. All these issues must be considered by every member involved in a service transaction in order to create a risk context and decide about proceeding with the transaction or not.

Finally, there is also a Historical Interactions category to consider those risk metrics related to the information and knowledge about the other entity involved in the transaction. In this phase, in contrast to the Pre-Federation case, there is another source to compute knowledge related metrics: the history of transactions. As more transactions are performed, the entities will have a better direct knowledge about the behavior of other entities. In addition, indirect knowledge could be obtained in anomalous situations. Consequently, the involved risk level can be tuned accordingly.

All the above distinctions are captured in the schematic of the taxonomy depicted in Figure 5.1.

Conclusions about the Taxonomy

To summarize, the taxonomy is organized into two major classes: Pre-Federation and Post-Federation. These taxonomic classes are further divided into the subclasses representing the following aspects where risk metrics must be placed: Security and Privacy, Knowledge, Interoperability, Service Specific and Historical Interactions. Finally, the last levels contemplate the different dimensions in which every risk can be evaluated. The taxonomy should be adopted by every entity in the system to enrich its intelligence and independence and to be capable of making well-informed decisions.

The classification compiles the characteristics of FIM systems and makes possible risk decomposition in small subsets. Besides, it is generic enough to be applicable to every federation framework (SAML, Liberty Alliance, OpenID, etc.), since the provided abstraction levels allow to cover all the common features, as well as the specific ones. Furthermore, exploiting the prediction capabilities of taxonomies, we observe a number of gaps that need to be covered to allow dynamic federation, namely:

1. The definition of standard formats to express security requirements/features.
2. The definition of standard formats to express legal, operational and technical features.
3. The incorporation of knowledge information to those frameworks that do not address this feature.
4. The introduction of mechanisms/formats to communicate this multi-protocol support for entities that support several identity protocols.

Finally, apart from serving as a basis for deriving risk metrics and identifying existing gaps, the taxonomy can be used to define an aggregation model for risk calculation. Its hierarchical structure makes it suitable to be the foundation of a hierarchical aggregation system. And this is an important advantage since multicriteria decision making (MCDM) [Triantaphyllou, 2000] mechanisms and related mathematical techniques rely on the decomposition of problems into a hierarchy of more easily comprehended sub-problems, each of which can be analyzed independently. We will show how to perform risk aggregation based on the taxonomy, but before this next section defines the metrics that will be used in our risk model.

5.3. Risk Metrics

5.3.1. What is a Metric?

"More than 100 years ago, Lord Kelvin insightfully observed that measurement is vital to deep knowledge and understanding in physical science. During the last few decades, researchers have made various attempts to develop measures and systems of measurement for computer security with varying degrees of success [Jansen, 2010]"

A metric is, according to the definition in [Jansen, 2010], a proposed measure or unit of measure that is designed to facilitate decision making and improve performance and accountability through collection analysis and reporting of relevant data. Another definition is that a metric is a measure for quantitatively assessing, controlling or selecting a person, process, event, or institution, along with the procedures to carry out measurements and the procedures for the interpretation of the assessment in the light of previous or comparable assessments[3]. In [Jaquith, 2007] metrics are defined as management tools that facilitate decision making and accountability through practical and relevant data collection, data analysis, and performance data reporting. As occurs with the concepts of trust, risk and reputation, the definition of metric is not universal and there are ambiguities and contradictions surrounding the term. In general all the definitions include the idea that metrics facilitate decision making, which is the main feature that motivates the use of metrics in the context of this dissertation. However, there are different opinions on whether the nature of a metric must be quantitative or qualitative and its implications. Some works, such as [Jaquith, 2007], suggest that qualitative assessments are bad metrics since they do not "count things" and so are subjective. It is true that, in general, quantitative metrics are more desirable than qualitative ones. However, it is challenging to find quantitative metrics that depict phenomena such as information security or too complex concepts such as trust. Furthermore, the distinction between quantitative and qualitative metrics can be easily obscured. For example, quantitative measures can be mapped to qualitative assignments and, in turn, numeric values can be used to represent rankings that are otherwise qualitative. In this regard, the terms measure and metric overlap in their definitions. Sometimes both concepts are referred as the same, but in other cases they are defined as not equal but related concepts. The NIST agency- in its project for software assurance SAMATE [Black, 2008]- makes the following differentiation between both terms: metrics are used for more abstract, higher-level, or somewhat subjective attributes, while measures are used for more concrete or objective attributes. But there are also contradictory opinions, such as the work in [Cugini et al., 1997] that defines metric and measure concepts in a totally opposite manner. Since the important point in describing a research work is to attach to a definition and be consistent, we will adopt the definitions in [Black, 2008]. Thus, we understand metrics as high level interpretations of objective measurements. Apart from the adoption of this definition we will focus on a particular kind of metrics: assurance metrics or levels of assurance. Assurance metrics represent objective confidence that an entity meets some requirements and they are based on specific evidence provided by the application of assurance techniques. The main feature of assurance metrics is that they provive certainty. As explained in [Vaughn Jr et al., 2003], the term assurance has been used for decades in trusted system development as an expression of confidence that one has in the strength of mechanisms or countermeasures.

[3] http://en.wikipedia.org/wiki/Metric_(unit)

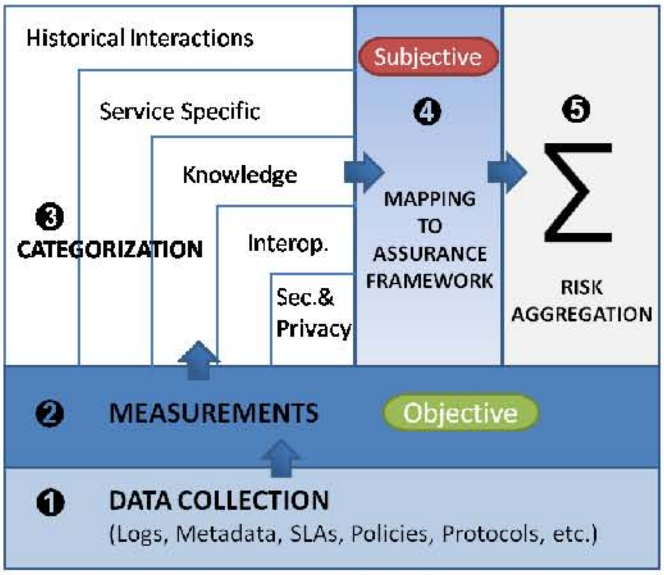


Figure 5.2: Conceptual description of the risk model and metrics

To conclude, our approach for risk assessment is based on assurance metrics, which are composed from low level measures. The measures are objective and built on well defined assurance descriptions, so different repetitions of the measurement procedure will lead to the same results. In turn, the assurance metrics are obtained from the combination of measures. Metrics are assigned a qualitative scale and thus are more subjective. However, the rationale behind the scale will be carefully explained. Finally, since is our goal to define a semiquantitative risk assessment framework, each qualitative category in the scale will be assigned to numerical values. Figure 5.2 shows the relationship between metrics and measures in our model and represents also the different metric categories that are grounded on the proposed taxonomy.

Next section goes through the details on how we define each metric. It starts with a review of current metrics used in FIM systems and then explains the metrics used in the proposed model.

5.3.2. Metrics for Risk Quantification in FIM

In our subject of study, risk assessment in FIM, there is still scarce work. Federated identity management has been researched during the last decade and it is becoming more and more important with the increasing segmentation on internet services and the penetration of the cloud computing paradigm. The need of sharing user identity data between online entities (e.g., attribute providers, banks, telcos and so on) is paramount for a flexible and seamless cooperation. But this sharing imposes risks; it is not the same relying in your own infrastructure for authenticating/verifying user's identity and give access to your services, than giving access by relying on what another entity claims about the user. This is were assurance comes into scene: it is crucial that the mechanisms used by cooperating partners are well-defined and meet each other's requirements, it is necessary to have a degree of certainty on which to build confidence.

Regarding metrics designed for risk assessment in FIM, there is important research on metrics for authentication assurance. More specifically, the identity assurance model developed by Kantara [Glade, 2012](originally started as the Liberty Alliance Identity Assurance Framework [Cutler, 2007]) constitutes one of the main assurance frameworks nowadays. Kantara defines the assurance levels that can be associated with a credential as measured by the associated technology, processes, and policy and practice statements. The Kantara framework defers to the guidance provided by the NIST Special Publication 800-63 [Nadalin et al., 2006], which outlines four levels of assurance, ranging in confidence level from low to very high. These four assurance levels, shown in Table 5.1, are used in relying parties to address increasing levels of risk, i.e., the choice of an assurance level is based on the degree of certainty of identity required to mitigate risk.

Level	**Description**
1	Little or no confidence in the asserted identity's validity
2	Some confidence in the asserted identity's validity
3	High confidence in the asserted identity's validity
4	Very high confidence in the asserted identity's validity

Table 5.1: Levels of authentication assurance as defined by Kantara[Glade, 2012]

Besides the work on authentication assurance carried out by Kantara, the ETSI also identifies metrics as a starting point for cooperation in federated environments [ETSI, 2011]. The ETSI vision is that there are several influences to decision making. These influences are in itself multi-factored and adequate metrics have to be developed to quantify and qualify them. Thus, apart from authentication assurance, they propose to include other metrics such as reputation, as it was explained in Chapter 2. Following the idea of using assurance metrics and inspired by the particular model of authentication assurance (the only well-defined assurance metric for FIM nowadays), we aim to design a comprehensive set of assurance metrics for the different risk categories identified in our taxonomy. We understand that assurance metrics directly affect the likelihood of attacks on related vulnerabilities. For example, a high level of confidentiality assurance will mean a low probability of successful attacks that compromise confidentiality, such as eavesdropping attacks. Thus, all the metrics used in our model are related to the likelihood part in the risk equation *Risk = Likelihood x Impact*. But in order to compute the risk, it is also required to evaluate the impact. For this purpose, we consider that impacts are related to each of the categories in the taxonomy, i.e., every assurance metric can be related to an associated impact. Since we aim to define a generic framework, we assume that impacts are qualified or described in a scale of values high, medium, and low, that is assigned by the entity evaluating the risk according to the value of its assets. These details will be furhter clarified at the end of the chapter, after all the assurance metrics are defined.

In the following we first introduce the procedure for metric identification based on the taxonomy. Next subsections will provide a description of the whole proposed set of basic metrics that are to be applied for risk assessment in FIM. In this regard, if an already existing metric fits in a taxonomic category, we will adopt this existing metric. However, for such cases where no existing metrics apply - which are the majority -, we will define a new metric. Finally, our framework is intended to be generic and thus usable for any FIM protocol. Based on this premise, the categories in the taxonomy cover general aspects involved in establishing and maintaining a federation independently of the underlying protocol. Nevertheless, the detailed definition of the metrics requires to be more specific on certain aspects, such as the identification of the data sources for measurements. Thus, whenever particularization is required, we will assume a SAML-based system.

Starting with the methodology to identify risk related metrics, the first step is to choose a terminal node in the taxonomy tree, e.g., *Post-Fed->Security and Privacy->Confidentiality->TL*. Then, for this selected category, the possible threats can be derived. In this example, if no confidentiality is provided at the transport level, the system could be subject of eavesdropping attacks or privacy violations. Consequently, transport confidentiality yields a contribution to the feasibility of the mentioned threats. However, the final risk will be affected by other components, such as the confidentiality at the message level. As previously mentioned, we rely on assurance metrics since they are inversely proportional to the feasibility of attacks related to the attribute being assured. Thus, instead of exhaustively composing a list of threats and computing the likelihood of each threat, we assume that the likelihood associated to an specific category generaly refers to the whole set of threats related to the category. According to this reasoning, table 5.2 shows the semantic definition of the metric. To allow quantification, a numeric value can be assigned depending on the cryptographic strength of the encryption algorithms supported by the entity willing to federate. We will later elaborate both on the qualitative and quantitative description of the metrics.

Metric	Confidentitality at transport level (CONF_{TL})
Definition	Measures confidentiality assurance of information exchanged at the transport level, depending on the encryption algorithm (e.g., based on strength, key size...)
Considerations	Source: can be obtained from the SAML entity metadata [Cantor et al., 2005c]

Table 5.2: Semantic definition of the confidentiality at transport level assurance metric

Following this strategy, we identify the rest of metrics. Another example is the level of authentication assurance or LOA (see Table 5.3). It is usually defined as the degree of confidence in identifying an entity to whom a credential was issued, and the degree of confidence that the entity using the credential is indeed the entity that the credential was issued to. For example, the NIST [Nadalin et al., 2006] agency defines four discrete levels that are associated with the strength of the authentication methods. This way, simple password challenge-response is categorized as LOA level 1, as well as hard cryptographic tokens are considered level 4. The LOA, which falls into the *Post-Fed->ServiceSpec* category, is a clear example of an existing concept that fits into the taxonomy and can be used as a metric for risk assessment in FIM. Thus, the value expressed by the LOA metric can be used by providers to decide whether an individual should be granted access to specific protected resources or if a higher LOA is required.

Metric	Level of authentication assurance (LOA)
Definition	Measures the degree of confidence in identifying an entity to whom a credential was issued
Considerations	Source: can be obtained from the SAML Authentication Assertion or from the metadata [Cantor et al., 2005c]

Table 5.3: Semantic description of the level of authentication assurance metric

Following the explained methodology we now derive assurance metrics for each of the categories in the taxonomy. For every metric we define its semantics, the procedures to carry out measurements and the associated qualitative scale. Summarizing, we answer these three questions:

- *What does the metric measure?*
- *How is the measure performed?*
- *What is the qualitative scale?*

Existing Applicable Metrics

As mentioned before, the only metrics whose usage is being adopted in the context of ad-hoc federation and trust management scenarios are levels of authentication assurance (LOA). More specifically, the NIST 800-63 definition, which is the reference followed by Kantara, is the most important framework.

1. What does the metric measure?

The LOA metric measures the degree of confidence in identifying an entity to whom a credential was issued, and the degree of confidence that the entity using the credential is indeed the entity that the credential was issued to.

2. How is the measure performed?

The NIST 800-63 definition provides a model for categorizing the wide variety of authentication methods into a small number of levels, typically based on some notion of the strength of the authentication. As described in [Nadalin et al., 2006], the procedure consists of evaluating the kind of tokens, credentials and protocols used for authentication, as well as the user registration process. To give an idea, Table 5.4 shows a classification of allowed token types at each LOA level. However, the set of features to be evaluated is far more complex. We do not include here all these details for simplicity, but we refer the interested reader to the original document.

In order to specify the source data that can be used in a FIM system to obtain the LOA metric following the measurement procedures, we particularize for the SAML case.

Token Type	Level 1	Level 2	Level 3	Level 4
Hard crypto token	Yes	Yes	Yes	Yes
One-time password device	Yes	Yes	Yes	Yes
Soft crypto token	Yes	Yes	No	No
Passwords & PINs	Yes	No	No	No

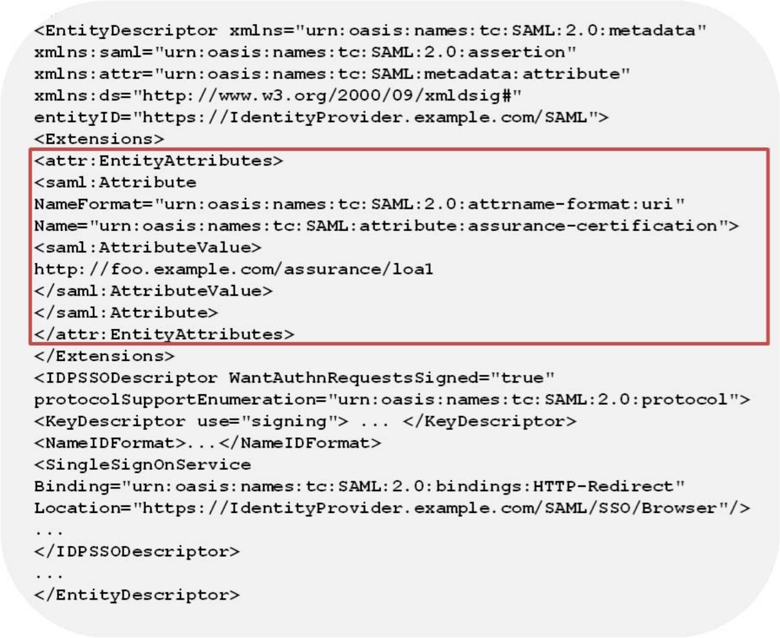
Table 5.4: Token types allowed at different LOA levels according to NIST SP 800-63 [Nadalin et al., 2006]

To derive the LOA in a SAML based system, entities should support the *"Level of Assurance Authentication Context Profiles for SAML 2.0"* [E. Tiffany, 2008]. This profile defines how to use existing SAML mechanisms to express identity assurance information, which can be done in two different manners: 1) using the SAML 2.0 Authentication Context [Kemp et al., 2005] mechanisms in order to allow SAML authentication requests and assertions to carry assurance information and 2) including extensions to SAML metadata documents [Cantor et al., 2005c] to represent assurance certification information about a SAML entity within the corresponding metadata.

The first mechanism is based on the definition of new SAML Authentication Context classes corresponding to different assurance criteria, thereby allowing the corresponding URIs for those assurance-based classes to be inserted within authentication requests and responses.

The second mechanism defines a SAML attribute profile that may be used to represent the certification status of an issuer of authentication statements (i.e., an Identity Provider) regarding its conformance with the requirements of an identity assurance framework. An example of a metadata document including this information is shown in Figure 5.3.

In the example in Figure 5.3 a `<saml:Attribute>` element is placed in the IdP's `<md:EntityDescriptor>` to indicate that the practices of the IdP have been certified as compliant with the requirements of level of authentication assurance 1 (LOA1) according to a fictional assurance framework (foo assurance framework) whose associated XML schema is located in the URL `http://foo.example.com/assurance/loa`. A party relying on this metadata could use this value as input for its risk policy and use it to decide whether to accept SAML authentication assertions from this IdP. Regarding the kind of criteria used to categorize each level of assurance, not only the NIST 800-63 framework is applicable. In fact, different definitions may be applied since the profile is not tied to a specific authentication assurance framework. Furthermore, there is an ongoing IETF draft



```
<EntityDescriptor xmlns="urn:oasis:names:tc:SAML:2.0:metadata"
xmlns:saml="urn:oasis:names:tc:SAML:2.0:assertion"
xmlns:attr="urn:oasis:names:tc:SAML:metadata:attribute"
xmlns:ds="http://www.w3.org/2000/09/xmldsig#"
entityID="https://IdentityProvider.example.com/SAML">
<Extensions>
<attr:EntityAttributes>
<saml:Attribute
NameFormat="urn:oasis:names:tc:SAML:2.0:attrname-format:uri"
Name="urn:oasis:names:tc:SAML:attribute:assurance-certification">
<saml:AttributeValue>
http://foo.example.com/assurance/loa1
</saml:AttributeValue>
</saml:Attribute>
</attr:EntityAttributes>
</Extensions>
<IDPSSODescriptor WantAuthnRequestsSigned="true"
protocolSupportEnumeration="urn:oasis:names:tc:SAML:2.0:protocol">
<KeyDescriptor use="signing"> ... </KeyDescriptor>
<NameIDFormat>...</NameIDFormat>
<SingleSignOnService
Binding="urn:oasis:names:tc:SAML:2.0:bindings:HTTP-Redirect"
Location="https://IdentityProvider.example.com/SAML/SSO/Browser"/>
...
</IDPSSODescriptor>
...
</EntityDescriptor>
```

Figure 5.3: Example of SAML metadata including LoA information (©[E. Tiffany, 2008])

aiming at establishing an IANA (Internet Assigned Numbers Authority) registry for LoA profiles [Johansson, 2012]. For example, apart from the NIST standard, there are other definitions of LoA such as the provided by the STORK Project [Clowes and Brathwaite, 2009].

3. What is the qualitative scale?

As stated, the qualitative definition depends on the LoA framework in use. For the case of using the NIST 800-63 definition, the qualitative scale is comprised of 4 levels, namely *Little* assurance, *Some* assurance, *High* assurance and *Very High* assurance.

In conclusion, Table 5.5 summarizes the complete definition of the LoA metric.

While the LOA metric is applicable for category *Post-Fed->ServiceSpec* in the taxonomy, no other well-defined assurance metrics were found that fit in the taxonomy. Thus, we proceed with the definition of new metrics. In this task, we will propose assurance frameworks to qualitatively evaluate each of the metrics. Whenever possible, the proposed framework will be based in widely accepted expert knowledge available in the related literature. For

[4] www.ref.gv.at/AG-IZ-Sicherheitsklassen-Sec.1719.0.html
[5] www.ref.gv.at/Sicherheitsklassen.2329.0.html
[6] www.eid-stork.eu

Metric	LOA
Definition	Level of authentication assurance. Measures the degree of confidence in identifying an entity to whom a credential was issued, and the degree of confidence that the entity using the credential is indeed the entity that the credential was issued to
Category	*Post-Fed->ServiceSpec*
Type	Basic
Formula	-
Measurement	The procedure consists of evaluating the kind of tokens, credentials and protocols used for authentication, as well as the user registration process. Once the strength is determined, the LOA is assigned to a qualitative category.
Source data	SAML metadata [Cantor et al., 2005c] or SAML Authentication Context [Kemp et al., 2005]
Qualitative Scale	▪ **Little confidence**, when the degree of confidence in the asserted identity's validity is little or nonexistent; ▪ **Some confidence**, when there is some confidence in the asserted identity's validity; ▪ **High confidence**, when the degree of confidence in the asserted identity's validity is high; ▪ **Very High confidence**, when the degree of confidence in the asserted identity's validity is very high
Applied Framework	NIST 800-63 [Nadalin et al., 2006]
Alternative Frameworks	ATSC2[4], ATSC3[5], STORK[6]

Table 5.5: LOA Metric Definition

those metrics that fall in a field where the available knowledge is scarce, a more high level framework will be proposed. In this section, only the basic metrics (i.e., those in the taxonomy leafs) are defined, since the intermediate metrics will be described later together with the aggregation procedure.

Proposal of New Metrics

Following the methodology previously explained, we use the taxonomy to identify the leaf nodes and define a set of basic metrics associated to them. These basic metrics will be later used as as basis to define aggregated metrics.

- **Pre-Fed->Security and Privacy->Confidentiality->TL**

  1. *What does the metric measure?*

  In this category we define the $\mathrm{CONF}_{\mathrm{TL}}$ metric, which measures confidentiality assurance of information exchanged at the transport level.

  2. *How is the measure performed?*

  The procedure for obtaining the confidentiality assurance categories at transport level consists of: (a) determining the algorithms/protocols used for information encryption; and (b) determining the assurance level according to the specific features of the encryption algorithms/protocols, (i.e., based on strength, key size, etc.) Regarding the kind of criteria used to categorize each level of assurance, we propose a simple framework, which is described in the following.

SAML specifications [Maler et al., 2005] recommend the use of the SSL/TLS protocol in order to protect the communications. The SSL/TLS server-client handshake involves negotiating cipher suites to be used for protecting the Internet transaction. A cipher suite combines four kinds of security features: 1) key establishment, 2) signature, 3) encryption algorithm, and 4) hash algorithm, and is given a name in the protocol specification. For example, the TLS_DHE_DSS_WITH_AES_256_CBC_SHA cipher suite uses:

1. The Ephemeral Diffie-Hellman key agreement, DHE,
2. The Digital Signature Standard, DSS (which implies the Digital Signature Algorithm, DSA),
3. The Asynchronous Encryption Algorithm AES with 256 bits key size in CBC (Cipher Block Chaining) mode , and,
4. The Secure Hash Algorithm, SHA-1 (used to compute a HMAC).

Here, since the purpose is to obtain a metric for confidentiality assurance, we start by providing a list of SSL/TLS cipher suites categorized according to the security of the encryption algorithm in use, which is the main feature related to confidentiality. These algorithms vary from very weak exportable ciphers such as RC4 in 40-bit mode to stronger ciphers such as 3DES or AES. Thus, as remarked by Rescorla in [Rescorla, 2001], *"It is therefore necessary to choose a cipher suite commensurate with the value of your data"*. For the elaboration of this list, depicted in Table 5.6, we use as knowledge base the NIST's document *"Guidelines for the Selection and Use of Transport Layer Security (TLS) Implementations"* [Chernick et al., 2005], the OWASP's *"Transport Layer Protection Cheat Sheet"* [7] and the SSL/TLS specifications [Dierks and Rescorla, 2006], [Dierks and Rescorla, 2008], [Hickman, 1995], [Freier et al., 2011].

[7] https://www.owasp.org/index.php/Transport_Layer_Protection_Cheat_Sheet

Cipher Suite	Encryption Algorithm	Key Size (bits)	Encryption Strength
TLS_DHE_DSS_WITH_AES_256_CBC_SHA TLS_DHE_RSA_WITH_AES_256_CBC_SHA TLS_RSA_WITH_AES_256_CBC_SHA TLS_DH_DSS_WITH_AES_256_CBC_SHA TLS_DH_RSA_WITH_AES_256_CBC_SHA TLS_DH_anon_WITH_AES_256_CBC_SHA TLS_RSA_WITH_AES_256_CBC_SHA256 TLS_DH_DSS_WITH_AES_256_CBC_SHA256 TLS_DH_RSA_WITH_AES_256_CBC_SHA256 TLS_DHE_DSS_WITH_AES_256_CBC_SHA256 TLS_DHE_RSA_WITH_AES_256_CBC_SHA256	AES_256_CBC	256	High
TLS_DHE_DSS_WITH_AES_128_CBC_SHA TLS_DHE_RSA_WITH_AES_128_CBC_SHA TLS_RSA_WITH_AES_128_CBC_SHA TLS_DH_DSS_WITH_AES_128_CBC_SHA TLS_DH_RSA_WITH_AES_128_CBC_SHA TLS_DH_anon_WITH_AES_128_CBC_SHA TLS_RSA_WITH_AES_128_CBC_SHA256 TLS_DHE_RSA_WITH_AES_128_CBC_SHA256 TLS_DHE_DSS_WITH_AES_128_CBC_SHA256 TLS_DH_DSS_WITH_AES_128_CBC_SHA256 TLS_DH_RSA_WITH_AES_128_CBC_SHA256	AES_128_CBC	128	High
TLS_DHE_DSS_WITH_3DES_EDE_CBC_SHA TLS_DHE_RSA_WITH_3DES_EDE_CBC_SHA TLS_RSA_WITH_3DES_EDE_CBC_SHA TLS_DH_DSS_WITH_3DES_EDE_CBC_SHA TLS_DH_RSA_WITH_3DES_EDE_CBC_SHA TLS_DH_anon_WITH_3DES_EDE_CBC_SHA TLS_KRB5_WITH_3DES_EDE_CBC_SHA TLS_KRB5_WITH_3DES_EDE_CBC_MD5 SSL_CK_DES_192_EDE3_CBC_WITH_MD5 SSL_RSA_WITH_3DES_EDE_CBC_SHA SSL_DH_DSS_WITH_3DES_EDE_CBC_SHA SSL_DH_RSA_WITH_3DES_EDE_CBC_SHA SSL_DHE_DSS_WITH_3DES_EDE_CBC_SHA SSL_DHE_RSA_WITH_3DES_EDE_CBC_SHA SSL_DH_anon_WITH_3DES_EDE_CBC_SHA	3DES_EDE_CBC, DES_192_EDE3_CBC	168	Medium
TLS_RSA_WITH_IDEA_CBC_SHA TLS_KRB5_WITH_IDEA_CBC_SHA TLS_KRB5_WITH_IDEA_CBC_MD5 SSL_CK_IDEA_128_CBC_WITH_MD5 SSL_RSA_WITH_IDEA_CBC_SHA	IDEA_CBC	128	Medium
SSL_CK_RC2_128_CBC_WITH_MD5	RC2_128	128	Medium

TLS_RSA_WITH_RC4_128_SHA TLS_RSA_WITH_RC4_128_MD5 TLS_DH_anon_WITH_RC4_128_MD5 TLS_KRB5_WITH_RC4_128_SHA TLS_KRB5_WITH_RC4_128_MD5 SSL_CK_RC4_128_WITH_MD5 SSL_RSA_WITH_RC4_128_MD5 SSL_RSA_WITH_RC4_128_SHA SSL_DH_anon_WITH_RC4_128_MD5 SSL_FORTEZZA_KEA_WITH_RC4_128_SHA	RC4_128	128	Medium
SSL_FORTEZZA_KEA_WITH_FORTEZZA_CBC_SHA SSL_FORTEZZA_KEA_WITH_RC4_128_SHA SSL_FORTEZZA_KEA_WITH_FORTEZZA_CBC_SHA	FORTEZZA_CBC	96	Medium
TLS_RSA_WITH_DES_CBC_SHA TLS_DH_DSS_WITH_DES_CBC_SHA TLS_DH_RSA_WITH_DES_CBC_SHA TLS_DHE_DSS_WITH_DES_CBC_SHA TLS_DHE_RSA_WITH_DES_CBC_SHA TLS_DH_anon_WITH_DES_CBC_SHA TLS_KRB5_WITH_DES_CBC_SHA TLS_KRB5_WITH_DES_CBC_MD5 SSL_CK_DES_64_CBC_WITH_MD5 SSL_RSA_WITH_DES_CBC_SHA SSL_DH_DSS_WITH_DES_CBC_SHA SSL_DH_RSA_WITH_DES_CBC_SHA SSL_DHE_DSS_WITH_DES_CBC_SHA SSL_DHE_RSA_WITH_DES_CBC_SHA SSL_DH_anon_WITH_DES_CBC_SHA	DES_CBC, DES_64_CBC	56	Low
TLS_RSA_EXPORT_WITH_DES40_CBC_SHA TLS_DH_DSS_EXPORT_WITH_DES40_CBC_SHA TLS_DH_RSA_EXPORT_WITH_DES40_CBC_SHA TLS_DHE_DSS_EXPORT_WITH_DES40_CBC_SHA TLS_DHE_RSA_EXPORT_WITH_DES40_CBC_SHA TLS_DH_anon_EXPORT_WITH_DES40_CBC_SHA TLS_KRB5_EXPORT_WITH_DES_CBC_40_SHA TLS_KRB5_EXPORT_WITH_DES_CBC_40_MD5 SSL_RSA_EXPORT_WITH_DES40_CBC_SHA SSL_DH_DSS_EXPORT_WITH_DES40_CBC_SHA SSL_DH_RSA_EXPORT_WITH_DES40_CBC_SHA SSL_DHE_DSS_EXPORT_WITH_DES40_CBC_SHA SSL_DHE_RSA_EXPORT_WITH_DES40_CBC_SHA SSL_DH_anon_EXPORT_WITH_DES40_CBC_SHA	DES_40_CBC	40	Low

TLS_RSA_EXPORT_WITH_RC2_CBC_40_MD5 TLS_KRB5_EXPORT_WITH_RC2_CBC_40_SHA TLS_KRB5_EXPORT_WITH_RC2_CBC_40_MD5 SSL_CK_RC2_128_CBC_EXPORT40_WITH_MD5 SSL_RSA_EXPORT_WITH_RC2_CBC_40_MD5	RC2_CBC_40	40	Low
TLS_RSA_EXPORT_WITH_RC4_40_MD5 TLS_DH_anon_EXPORT_WITH_RC4_40_MD5 TLS_KRB5_EXPORT_WITH_RC4_40_SHA TLS_KRB5_EXPORT_WITH_RC4_40_MD5 SSL_CK_RC4_128_EXPORT40_WITH_MD5 SSL_RSA_EXPORT_WITH_RC4_40_MD5 SSL_DH_anon_EXPORT_WITH_RC4_40_MD5	RC4_40	40	Low
TLS_NULL_WITH_NULL_NULL TLS_RSA_WITH_NULL_MD5 TLS_RSA_WITH_NULL_SHA TLS_RSA_WITH_NULL_SHA256 SSL_FORTEZZA_KEA_WITH_NULL_SHA SSL_NULL_WITH_NULL_NULL SSL_RSA_WITH_NULL_MD5 SSL_RSA_WITH_NULL_SHA SSL_FORTEZZA_KEA_WITH_NULL_SHA	NULL (No encryption)		

Table 5.6: Encryption strength provided by algorithms in SSL/TLS cipher suites

To break a communication session, an attacker can attempt to break the symmetric cipher used for the bulk of the communication. So it is obvious that a stronger cipher allows for stronger encryption and thus increases the effort needed to break it. For this reason it is recommended to use strong ciphers with sufficiently large key sizes and the classification we make in Table 5.6 is to be used for that purpose. In order to decide on classification into *High*, *Medium* and *Low* categories, we have followed these arguments:

- Algorithms with small key sizes (40, 56 bits) and DES are considered to provide low encryption strength. The reason is that they are identified as deprecated in [Chernick et al., 2005] and also these algorithms are not recommended as a good practice in many security guides in the literature [Smart, 2010].
- Algorithms AES and 3DES are considered to provide high encryption strength. The reason is that they are recommended as the most secure options in [Cher-

nick et al., 2005].

- For the rest of algorithms that cannot be directly categorized as having low or high strength according to the documentation used, we decided to apply the key-length criterion: the longer the key, the more secure. Based on this, those algorithms with a key size smaller than 128 are considered to provide low encryption strength; and those whose key size is equal to 128, are considered to provide medium encryption strength. .

It is to note that, despite there are around 200 known cipher suites, Table 5.6 includes only those listed in the SSL/TLS RFCs and specifications. Thus, encryption algorithms such as for example SEED or Camellia, are not considered in this classification despite the existence of cipher suites including them. However, we believe that the analyzed subset of cipher suites is enough for federation scenarios, since it covers the most frequently used algorithms and those used and recommended by FIM specifications.

Furthermore, apart from the encryption algorithm, the protocol version has also an impact in the feasibility of confidentiality-related attacks. Despite that the terms Secure Socket Layer (SSL) and Transport Layer Security (TLS) are often used interchangeably, different versions of SSL and TLS are supported by clients and servers. Weaknesses have been identified in SSL 2.0 and SSL 3.0 and they have been successfully attacked, which makes these versions insecure. The best practice recommended for transport layer protection is to only provide support for the TLS protocols - TLS 1.0, TLS 1.1 and TLS 1.2. In situations where lesser security requirements are necessary, it may be acceptable to also provide support for SSL 3.0. But in no situation should SSL 2.0 be enabled , since its weaknesses are such that the provided transport layer protection is not effective. According to that recommendations, we apply also these additional rules for categorization:

- If the protocol in use is SSL 2.0, the encryption strength is penalized and decreased to Low encryption strength.
- If the protocol in use is SSL 3.0, the encryption strength is penalized by decreasing the category to the immediate lower level (except in the case of low strength, which is maintained as low since there is no lower level).

In order to include these new rules in the construction of the confidentiality assurance metric, Table 5.7 shows the Protocol Penalized Encryption Strengths (PPES) according to the combination of encryption algorithm and protocol.

Encryption Strength	**Protocol**	**PPES**
High	TLS	High
	SSL 3.0	Medium
	SSL 2.0	Low
Medium	TLS	Medium
	SSL 3.0	Low
	SSL 2.0	Low
Low	TLS	Low
	SSL 3.0	Low
	SSL 2.0	Low

Table 5.7: Protocol Penalized Encryption Strength (PPES) provided by SSL/TLS cipher suites

As the last step to define the confidentiality assurance metric, another part of the SSL/TLS protocol must be considered due to its impact in confidentiality: the key establishment procedure. Key establishment is the process of establishing a shared secret key (or the material to derive this key) that will be used for subsequent cryptographic operations over a SSL/TLS connection (i.e., for encryption and hashing). The key establishment algorithms used in SSL/TLS [Chernick et al., 2005] are 1) RSA, 2) Diffie-Hellman in three possible variants (static, ephemeral, anonymous) and 3) Fortezza-KEA. Options RSA, Fortezza-KEA and the static and ephemeral variants of DH imply authenticated exchange; while anonymous DH means that no authentication is performed during the exchange. We refer the interested reader to [Chernick et al., 2005] for more details about these mechanisms, but the goal is basically ensuring the safe generation and exchange of the secret keys that will be used during the remainder of the session. Thus, weaknesses in the key exchange phase can lead to man-in-the-middle (MITM) attacks that allow the attacker gaining access to the complete communication channel or make the per-session secret keys easier to compromise. For these reasons, we apply a final refinement to define the confidentiality assurance metric in Table 5.8. We divide the key establishment procedures into authenticated with ephemeral parameters, authenticated with non-ephemeral parameters and anonymous (i.e., non-authenticated). Then, we construct our metric by applying the following rules:

- If an authenticated key establishment procedure with ephemeral parameters is used, the confidentiality strength is maintained as is. The reason underlying this rule is that [Chernick et al., 2005] recommends RSA or DSA authentication with ephemeral Diffie-Hellman key agreement for maximum security, as it allows perfect forward secrecy.

- If an authenticated key establishment procedure with non-ephemeral parameters is used, the confidentiality strength is penalized by decreasing its level to the immediate lower level category (except in the case of low, which is maintained as is since there is no lower level). The reason underlying this rule is that, as derived from [Chernick et al., 2005], authentication key exchange procedures with non ephemeral parameters are less secure than ephemeral ones since they do not provide perfect forward secrecy.

- If an anonymous key establishment procedure is used (i.e., no authentication), then the penalized confidentiality strength is decreased to the low level. The reason underlying this rule is that anonymous key exchange mechanisms are not recommended because they are subject to MITM attacks. As explained in [Dierks and Rescorla, 2006], completely anonymous connections only provide protection against passive eavesdropping. Unless an independent tamper-proof channel is used to verify that the finished messages are not replaced by an attacker, server authentication is required in environments where active man-in-the-middle attacks are a concern.

PPES	**Kind of Key Establishment**	**Confidentiality Assurance -Transport Level ($\mathbf{CONF_{TL}}$)**
High	Authenticated Ephemeral	High
	Authenticated Non-Ephemeral	Medium
	Anonymous	Low
Medium	Authenticated Ephemeral	Medium
	Authenticated Non-Ephemeral	Low
	Anonymous	Low
Low	Authenticated Ephemeral	Low
	Authenticated Non-Ephemeral	Low
	Anonymous	Low

Table 5.8: Confidentiality Assurance Metric provided by SSL/TLS cipher suites based on PPES and Key Establishment algorithms

Of course the kind of authenticated mechanism will make the exchange procedure more or less secure, but this aspect will be measured by the authentication assurance metric defined later in this section.

Finally, another common option available to secure an Internet connection is the creation of Virtual Private Networks (VPN) using the IPSec protocol. We have focused in SSL/TLS because it is the model recommended by SAML (and the rest of FIM standards) and the most commonly used in the scenarios that are object of this thesis.

The source data to evaluate $\mathrm{CONF}_{\mathrm{TL}}$ assuming a SAML-based system can be the metadata. There is a recent specification, entitled *"SAMLv2.0 Metadata Profile for algorithm support"* [Cantor, 2010], which defines the extensions to convey this kind of cryptographic information. The problem is that, so far, the specification is intended to inform about the algorithms used at SAML message level. In our opinion it would be reasonable and useful to add support to inform about the available cipher suites at transport level, whether in the metadata or in specific security SLAs.

3. What is the qualitative scale?

The qualitative scale used for the confidentiality assurance at transport level metric encompasses three levels: *Low*, *Medium*, and *High*. Each of these levels represents the assurance achieved in confidentiality depending on the SSL/TLS cipher suite in use, and according to criteria 1) encryption strength, 2) protocol version and 3) key establishment procedure.

- **Pre-Fed->Security and Privacy->Confidentiality->ML**

1. What does the metric measure?

In this category we define the $\mathrm{CONF}_{\mathrm{ML}}$ metric, which measures confidentiality assurance of information exchanged at the message level .

2. How is the measure performed?

The procedure for obtaining the assurance categories encompasses the same steps described for the $\mathrm{CONF}_{\mathrm{TL}}$ metric, but taking into account the algorithms used at the message level. Regarding the kind of criteria used to categorize each level of assurance, we propose a simple framework, which is explained in the following.

SAML protocol messages can be protected in regard to confidentiality by means of XMLEncryption [Eastlake et al., 2002b]. As in the case of confidentiality at transport level, the degree of protection assurance will depend on the strength of the encryption algorithms used. The XMLEncryption specification defines a set of algorithms and the associated requirements for implementation. More specifically, for the block ciphers category it requires implementation of AES with key sizes 128 and 256 bits, and Triple DES; and recommends AES with 192 key size as optional implementation. On the other hand, no specific stream encryption algorithms are specified. Furthermore, the mechanism is extensible, and alternative algorithms may be used- though no extensions have been defined to date. Based on this, we define a simple framework for confidentiality assurance at message level, which is described in Table 5.9.

Encryption algorithm	**Confidentiality Assurance - Message level (CONF$_{\mathrm{ML}}$)**
AES_CBC_128 AES_CBC_192 AES_CBC_256 3DES_EDE_CBC	High

Table 5.9: Confidentiality assurance at message level provided by encryption algorithms in XMLEncryption

It is to note that the XMLEncryption standard was published after algorithms like DES or those with small key sizes (80 bits and lower) had been declared as deprecated due to well-known weaknesses. For this reason, only strong algorithms are included as supported by the specification. Due to this fact, we only consider a single category for the metric that maps these strong algorithms to a *High* confidentiality assurance.

The source data to evaluate CONF$_{\mathrm{ML}}$ assuming a SAML-based system can be the metadata, where supported cryptographic algorithms can be communicated according to the extensions defined in [Cantor, 2010].

3. What is the qualitative scale?

As in the case of the confidentiality metric at transport level, the qualitative scale we aim to use for the confidentiality assurance at message level is the three level scale *Low*, *Medium*, and *High*. Each of these levels represents the assurance achieved in confidentiality depending on the strength of the encryption algorithm in use. Despite currently there are only high assurance algorithms, the framework will evolve when new algorithms appear and the current ones become less secure, and so the different scale levels may have an algorithm mapping.

- **Pre-Fed->Security and Privacy->Integrity->TL**

1. What does the metric measure?

In this category we define the $\mathrm{INT}_{\mathrm{TL}}$ metric, which measures integrity assurance of information exchanged at the transport level.

2. How is the measure performed?

The procedure for obtaining the integrity assurance categories at transport level consists of: (a) determining the algorithms/protocols used for integrity protection; and (b) determining the assurance level according to the specific features of these algorithms/protocols. Regarding the kind of criteria used to categorize each level of assurance, we propose a simple framework, which is described in the following.

As previously mentioned, SAML specifications [Maler et al., 2005] recommend the use of the SSL/TLS protocol for protection of the communications at transport level. Now we analyze the available SSL/TLS cipher suites regarding their integrity capabilities. In these sense the hash algorithm in use determines the integrity assurance. Therefore, with the purpose of defining a metric for integrity assurance, we perform an analysis similar to the employed in the definition of the confidentiality at transport level metric. First of all, the hash algorithms used in SSL/TLS cipher suites are categorized according to the security provided. For the elaboration of this list we use as knowledge base the same documents used for defining the confidentiality metric at transport level, plus and an additional NIST's document entitled *"Recommendation for Applications Using Approved Hash Algorithms"* [Dang, 2008]. In this case, for the sake of better readability, we have simplified the classification (see Table 5.10) by eliminating the cipher suite names, since they were already listed in Table 5.6.

Hash Algorithm	Digest Length (bits)	Integrity Strength
SHA256	256	High
SHA	160	Medium
MD5	128	Low
NULL	-	No Integrity

Table 5.10: Integrity strength provided by hash algorithms used SSL/TLS cipher suites

As shown in Table 5.10, SSL/TLS offers two options for a cryptographic hash algorithm: SHA (with different digest sizes) and MD5. In order to decide on classification into *High*, *Medium* and *Low* categories, we have followed these arguments:

- The MD5 algorithm is considered to provide low integrity. The reason is that this algorithm is known to be weak and it is not recommended in [Chernick et al., 2005].
- The SHA algorithm is considered to provide higher levels of integrity assurance than MD5, since it is recommended as a better option in [Chernick et al., 2005]. Furthermore, since the strength of hash algorithms is higher when the size of the digest is longer [Dang, 2008], SHA256 is considered to provide high assurance, and SHA is considered to provide medium assurance.

Furthermore, there are two more factors that affect the integrity metrics. Firstly, the SSL/TLS protocol version in use. And secondly the key establishment process, since integrity is based on the combination of the hash function with a key (i.e., in the HMAC code) that is set in this establishment phase. Thus, in order to include the impact of both factors in the final integrity assurance, we follow the same rules used in the definition of the CONF_{TL} metric.

Accordingly, Table 5.11 shows the Protocol Penalized Integrity Strength (PPIS), and

Integrity Strength	Protocol	PPIS
High	TLS	High
	SSL 3.0	Medium
	SSL 3.0	Low
Medium	TLS	Medium
	SSL 3.0	Low
	SSL 3.0	Low
Low	TLS	Low
	SSL 3.0	Low
	SSL 3.0	Low

Table 5.11: Protocol Penalized Integrity Strength (PPIS) provided by SSL/TLS cipher suites

Table 5.12 summarizes the final metric for integrity assurance at transport level.

PPIS	Kind of Key Establishment	Integrity Assurance -Transport Level ($\mathbf{INT_{TL}}$)
High	Authenticated Ephemeral	High
	Authenticated Non-Ephemeral	Medium
	Anonymous	Low
Medium	Authenticated Ephemeral	Medium
	Authenticated Non-Ephemeral	Medium
	Anonymous	Low
Low	Authenticated Ephemeral	Low
	Authenticated Non-Ephemeral	Low
	Anonymous	Low

Table 5.12: Integrity Assurance Metric provided by SSL/TLS cipher suites based on PPIS and Key Establishment algorithms

As in the case of the $\mathrm{CONF_{TL}}$ metric, the kind of authenticated mechanism will also make the exchange procedure more or less secure, but this aspect will be measured by the authentication assurance metric defined later in this section. Regarding the source data to evaluate $\mathrm{INT_{TL}}$ assuming a SAML-based system, the information could be extracted from the metadata or from the SLAs, if appropriate mechanisms to include this information are defined.

3. What is the qualitative scale?

The qualitative scale used for the integrity assurance at transport level metric encompasses three levels: *Low*, *Medium*, and *High*. Each of these levels represents the assurance achieved in integrity depending on the SSL/TLS cipher suite in use, and according to criteria 1) integrity strength, 2) protocol version and 3) key establishment procedure.

- **Pre-Fed->Security and Privacy->Integrity->ML**

1. What does the metric measure?

In this category we define the $\mathrm{INT_{ML}}$ metric, which measures integrity assurance of information exchanged at the message level.

2. How is the measure performed?

The procedure for obtaining the integrity assurance categories encompasses the same steps described for the $\mathrm{INT_{TL}}$ metric, but taking into account the algorithms used at the message level. Regarding the kind of criteria used to categorize each level of assurance, we propose a simple framework, which is explained in the following.

SAML specifications contemplate that application messages can be protected in regard to integrity by means of XMLSignature [Eastlake et al., 2002a] [Eastlake et al., 2008], [Eastlake et al., 2012]. As in the case of integrity at transport level, the degree of protection assurance will depend on the strength of the digest algorithms in use. In this case, the digest is used in conjunction with a signature algorithm, so the strength of this signature algorithm will also have an impact on the final metric. On the one hand , the XMLSignature specifications contemplate the following digest algorithms: MD5, SHA, SHA-256, SHA-384 y SHA-512. On the other hand, the available signature algorithms are RSA, DSA and ECDSA (Elliptic Curve DSA). Based on this, we define a simple framework for integrity assurance at message level, which is described in Table 5.13.

In order to decide on classification into *High*, *Medium* and *Low* categories, we have followed these rules:

- First, we categorize the strength of the digest algorithms into low, medium and high levels by applying the same arguments as for the $\mathrm{INT_{TL}}$ metric.
- Next, we categorize the strength of the signature algorithms based on the key size: the stronger the key, the more secure is the signature. For this classification, we base on the framework for comparable key sizes in different algorithms provided in [Barker et al., 2011] and follow their recommendations by applying the following rules:
    - Signatures performed using RSA/DSA with key sizes greater than or equal to 3072 bits; or using ECDSA with key size greater than or equal to 256, are considered as high strength signatures.
    - Signatures performed using RSA/DSA with key sizes smaller than 3072 and greater or equal to 2048 bits; or using ECDSA with key size smaller than

Hash Algorithm	**Digest Length (bits)**	**Digest Strength**	**Signature Algorithm -Key size (bits)**	**Integrity Assurance - Message level (INT_{ML})**
SHA256, SHA384, SHA512	256, 384, 512	High	RSA/DSA $\geq$ 3072, ECDSA $\geq$ 256	High
			RSA/DSA $\geq$ 2048 and $<$ 3072, ECDSA $\geq$ 224 and $<$ 256	Medium
			RSA/DSA <2048, ECDSA <224	Low
SHA	160	Medium	RSA/DSA $\geq$ 3072, ECDSA $\geq$ 256	Medium
			RSA/DSA $\geq$ 2048 and $<$ 3072, ECDSA $\geq$ 224 and $<$ 256	Medium
			RSA/DSA <2048, ECDSA <224	Low
MD5	128	Low	RSA/DSA $\geq$ 3072, ECDSA $\geq$ 256	
			RSA/DSA $\geq$ 2048 and $<$ 3072, ECDSA $\geq$ 224 and $<$ 256	Medium
			RSA/DSA <2048, ECDSA <224	Low

Table 5.13: Integrity assurance at message level provided by XMLSignature hash and signature algorithms

256 and greater or equal to 224 bits, are considered as medium strength signatures.

- Signatures performed using RSA/DSA with key sizes smaller than 2048 bits; or using ECDSA with key size smaller than 256, are considered as low strength signatures.

- Finally, the digest strength is combined with the signature strength by choosing the lowest value (e.g., if the digest strength is high and the key length is low, the integrity assurance will be low).

The source data to evaluate INT_{ML} assuming a SAML-based system can be the metadata, where supported cryptographic algorithms can be communicated according to the extensions defined in [Cantor, 2010].

3. What is the qualitative scale?

As in the case of the metrics defined so far, we adopt a three level qualitative scale that encompasses *Low*, *Medium*, and *High* levels. Each of these levels represents the assurance achieved in integrity depending on the strength of XMLSignature hash and signature algorithms in use.

- **Pre-Fed->Security and Privacy->Authentication->TL**

1. What does the metric measure?

In this category we define the $\mathrm{AUTH}_{\mathrm{TL}}$ metric, which measures authentication assurance at the transport level.

2. How is the measure performed?

The procedure for obtaining the authentication assurance categories at transport level consists of: (a) determining the mechanisms used for authentication; and (b) determining the assurance level according to the specific features of these mechanisms. Regarding the kind of criteria used to categorize each level of assurance, we propose a simple framework, which is described in the following.

Based on the recommendation of using the SSL/TLS protocol for protection of the communications at transport level in SAML, we analyze now the authentication capabilities offered by the protocol. In this regard, the protocol aspect related with authentication is the key exchange procedure. This process serves two functions: 1) ensuring the safe generation and exchange of the secret keys that will be used during the remainder of the session; and 2) perform authentication, allowing at least one party to verify the identity of the other party. Function 1) affects both integrity and confidentiality of the session and thus was considered in the definition of the $\mathrm{CONF}_{\mathrm{TL}}$ and $\mathrm{INT}_{\mathrm{TL}}$ metrics. Function 2) is analyzed here with the aim to to define our authentication assurance metric at transport level.

In order to perform authentication, an entity involved in an SSL/TLS dialogue presents its digitally signed certificate to the other party and additionally the entity may also sign some data or use public cryptography. In either case, the client can verify the certificate or signature to ensure that the parameters belong to the server. Thus, presenting a valid certificate and proving possession of the private key authenticates the presenter to the recipient. Based on the same documents used to define the CONF$_{\text{TL}}$, we define the authentication assurance metric at transport level as depicted in Table 5.14.

Key Exchange Mechanism	**Description**	**Key Exchange Strength**	**Certificate Key Lengt (bits)**	**Authentication Assurance - Transport level (AUTH$_{\text{TL}}$)**
DHE_DSS	Ephemeral DH with DSS signatures	High	≥ 3072	High
			$\geq 2048, < 3072$	Medium
			<2048	Low
DHE_RSA	Ephemeral DH with RSA signatures	High	≥ 3072	High
			$\geq 2048, < 3072$	Medium
			<2048	Low
RSA	RSA key exchange, RSA certificate and public cryptography	High	≥ 3072	Medium
			$\geq 2048, < 3072$	Medium
			<2048	Low
DH_DSS	DH with DSS-based certificates	Medium	≥ 3072	Medium
			$\geq 2048, < 3072$	Medium
			<2048	Low
DH_RSA	DH with RSA-based certificates	Medium	≥ 3072	Medium
			$\geq 2048, < 3072$	Medium
			<2048	Low
Fortezza_KEA	Key Exchange Algorithm (KEA), DSS signature	Low	≥ 3072	Low
			$\geq 2048, < 3072$	Low
			<2048	Low
DH_anon	Anonymous DH, no certificate, no signatures			No assurance

Table 5.14: Authentication assurance metric at transport level provided by SSL/TLS cipher suites

In order to decide on classification into *High*, *Medium* and *Low* categories for the authentication metric, we define a framework based on the following rules:

- Firstly, the strength of the key exchange algorithm in regard to authentication is rated. We consider as high strength algorithms those using an authenticated method including signatures or public key cryptography. We consider as medium strength algorithms those using authenticated methods with certificates. Justification for this categorization lies on the fact that adding signatures or cryptography provides an additional degree of authentication apart from just presenting the certificate. In regard to the signature algorithm we make no further distinction because there are no proofs that choosing DSA over RSA or viceversa is better from a security perspective. Furthermore, Fortezza is considered as a low strength algorithm because the IETF standards committee did not include it in TLS 1.0. And, finally, anonymous algorithms are directly considered as providing no authentication assurance since they do not apply any mechanism for this purpose.

- Then, the certificate key length is considered as a parameter to refine the categorization and to obtain the final metric. This is because the stronger the key, the more difficult it is to break the key exchange phase and more secure is the authentication. Based on the recommendation in [Barker et al., 2011] that security applications should use "*at least 2048-bit public keys for securing information beyond 2010 (and 3072-bit keys for securing information beyond 2030)*" we consider:
    - Key sizes greater than or equal to 3072 bits as high strength keys,
    - Key sizes smaller than 3072 and greater or equal to 2048 bits as medium strength keys, and
    - Key sizes smaller than 2048 bits as mow strength keys
- Finally, the key strength is combined with the key exchange strength by choosing the lowest value (e.g., if the key exchange is high and the key length is low, the Authentication assurance will be low).

For the classification it is assumed that the certificate and/or the signatures are valid. Also, for the sake of better readability the classification in table 5.14 does not include the cipher suite names, since they were already listed in Table 5.6.

Regarding the source data to evaluate AUTH_{TL} assuming a SAML-based system, the information could be extracted from the metadata or from the SLAs, if appropriate mechanisms to include this information are defined.

3. What is the qualitative scale?

As in the case of the metrics defined so far, we adopt a three level qualitative scale that encompasses *Low*, *Medium*, and *High* levels. Each of these levels represents the assurance achieved in authentication depending on the features of the authentication mechanisms in use.

- **Pre-Fed->Security and Privacy->Authentication->ML**

1. What does the metric measure?

In this category we define the AUTH$_{\mathrm{ML}}$ metric, which measures authentication assurance of information exchanged at the message level.

2. How is the measure performed?

The procedure for obtaining the authentication assurance categories encompasses the same steps described for the AUTH$_{\mathrm{TL}}$ metric, but taking into account the algorithms used at the message level. Regarding the kind of criteria used to categorize each level of assurance, we propose a simple framework, which is explained in the following.

SAML specifications contemplate that application messages can be protected in regard to authentication by means of XMLSignature [Eastlake et al., 2002a] [Eastlake et al., 2008], [Eastlake et al., 2012].

To obtain the authentication assurance in this case we consider the signature options available in XMLSignature. The specifications contemplate the following signature algorithms: RSA, DSA and ECDSA (Elliptic Curve DSA). The security level associated to the algorithm basically depends on the size of the keys used. Based on this, we define a simple framework for authentication assurance at message level, which is described in Table 5.15.

Signature Algorithm -Key size (bits)	**Authentication Assurance - Message level (AUTH$_{\mathrm{TL}}$)**
RSA/DSA $\geq$ 3072, ECDSA $\geq$ 256	High
RSA/DSA $\geq$ 2048 and $<$ 3072, ECDSA $\geq$ 224 and $<$ 256	Medium
RSA/DSA <2048, ECDSA <224	Low

Table 5.15: Authentication assurance at message level provided by XMLSignature signature algorithms

The categorization into *High*, *Medium* and *Low* level is performed as done for the INT$_{\mathrm{ML}}$ metric based on the recommendations in [Barker et al., 2011].

Regarding the source data to evaluate AUTH_{ML} assuming a SAML-based system, the information could be extracted from the metadata, where supported cryptographic algorithms can be communicated according to the extensions defined in [Cantor, 2010].

3. What is the qualitative scale?

As in the case of the metrics defined so far, we adopt a three level qualitative scale that encompasses *Low*, *Medium*, and *High* levels. Each of these levels represents the assurance achieved in authentication depending on the strength of the XMLSignature signature algorithm in use.

- **Pre-Fed->Security and Privacy->Non Repudiation**

1. What does the metric measure?

In this category we define the NON-REP metric, which measures the degree of non-repudiation assurance.

2. How is the measure performed?

The procedure for obtaining the non-repudiation assurance categories consist of: (a) determining the mechanisms used to achieve non-repudiation; and (b) determining the assurance level according to the specific features of these mechanisms. Regarding the kind of criteria used to categorize each level of assurance, we propose a simple framework, which is described in the following.

Non-repudiation is generally implemented through the use of digital signatures, but more complex protocols and mechanisms have been developed that provide higher levels of assurance [Kremer et al., 2002]. In the case of SAML, the use of XMLSignature is recommended to sign the exchanged messages, but no additional mechanisms are sugested to improve non-repudiation. With the aim to contemplate both features, we propose the assurance levels summarized in Table 5.16.

The classification into *High*, *Medium* and Low levels is based on the fact that signatures alone provide less non-repudiation guarantees than if additional techniques are applied. And, in the case of using digital signatures, the key size makes a difference

Mechanism description	**Non-repudiation assurance (NON-REP)**
Usage of signatures plus additional non-repudiation mechanisms	High
Usage of signatures with key sizes RSA/DSA $\geq$ 2048, ECDSA $\geq$ 224	Medium
Usage of signatures with key sizes RSA/DSA <2048 bits, ECDSA <224 bits	Low

Table 5.16: Non-repudiation assurance framework

in the the security strength, as analyzed before.

As source data to measure non-repudiation assuming a SAML based system, metadata and SLAs can be used.

3. What is the qualitative scale?

As in the case of the metrics defined so far, we adopt a three level qualitative scale that encompasses *Low*, *Medium*, and *High* levels. Each of these levels represents the assurance achieved in non-repudiation depending on the strength of the XMLSignature signature algorithm in use.

- **Pre-Fed->Security and Privacy->Availability**

1. What does the metric measure?

In this category we define the AV metric, which measures the degree of availability assurance.

2. How is the measure performed?

The system availability is a well defined value that represents the average percentage of time a service is available. The availability is in fact a typical security feature included and well defined in SLAs. Based on this, the procedure for obtaining the availability assurance provided by a system consists of gathering the availability value and determining its associated assurance category. Regarding the kind of criteria used to categorize each level of assurance, we propose a simple framework, which is depicted in Table 5.17.

SAML specifications do not include recommendations on availability, neither provide a standardized way to convey this information. For these reasons, we define a simple

Average availability (aav) range	**Availability Assurance (AV)**
aav $\geq$ 99%	High
$\geq$ 97% aav < 99%	Medium
$\geq$ 95% aav <97%	Low

Table 5.17: Availability assurance framework

high-level framework based on the categorization of availability values in ranges.

The source data could be the SLAs, if appropriate mechanisms to include this information are defined.

3. What is the qualitative scale?

As in the case of the metrics defined so far, we adopt a three level qualitative scale that encompasses *Low*, *Medium*, and *High* levels. Each of these levels represents the assurance achieved in availability depending on the value of the average percentage of time the service is available.

- **Pre-Fed->Security and Privacy->Accountability**

1. What does the metric measure?

In this category we define the ACC metric, which measures the degree of accountability assurance.

2. How is the measure performed?

Since accountability is defined as the ability to associate a consequence with a past action of an individual, information logging is required for an entity to be accountable. Depending on the information stored in these logs, on the time this information is stored and on a number of other features, the accountability strength will be different. Based on this, the procedure for obtaining the accountability assurance provided by an entity consists of evaluating the accountability requirements fulfilled by the entity and determining its associated assurance category. Currently, there is not a well defined list of requirements, neither recommendations that can be used as objective criteria to elaborate a detailed framework. Thus, we propose a simple high-level framework, which is depicted in Table 5.18.

As it can be seen, the framework definition provided in this case abstracts the un-

Accountability Assurance (ACC)	**Description**
High	If a range of 70% 100% of optional accountability requirements is fulfilled apart from the mandatory ones
Medium	If a range of 35% 70% of optional accountability requirements is fulfilled apart from the mandatory ones
Low	If the set of minimum accountability requirements is fulfilled

Table 5.18: Accountability assurance framework

derlying criteria. It is assumed that a list of standard requirements is provided that can be checked against a local policy to determine the assurance. In general, there is a lack of applicable standards at this time, so providers and customers must work together to determine the information needed and how to make it available. In the particular case of SAML, we have reviewed publicly available policies from federations deployed in the real world [TERENA, 2012] and the requirements with regard to accountability are variable. For example, some federations require that specific information is logged by the different entities (e.g., username, timestamp, etc.) and retained for a specified period of time; while other federations do not impose requirements on auditing or let the involved entities free to choose and apply their own policies. Thus, it is clear that an accountability approach requires organizations to establish policies consistent with recognized external criteria. Due to the lack of information and the scarce work on the field, the identification of a set of standard requirements for FIM accountability and the analysis of their strength will be proposed as a future research line.

Assuming a SAML-based system, the source data to obtain the accountability metric can be the metadata, SLAs and policies if appropriate mechanisms are defined. For example, the UK Federation[8] policy documents define a mechanism for using, metadata to indicate the support of user accountability.

3. What is the qualitative scale?

As in the case of the metrics defined so far, we adopt a three level qualitative scale that encompasses *Low*, *Medium*, and *High* levels. Each of these levels represents the assurance achieved in accountability depending on the strength of the accountability policies in use.

- **Pre-Fed->Security and Privacy->Privacy** *1. What does the metric measure?*

[8] http://www.ukfederation.org.uk/library/uploads/Documents/federation-technical-specifications.pdf

In this category we define the LOP metric, which measures the Level of Protection assurance, i.e., the degree of data protection in regard to privacy.

2. How is the measure performed?

The procedure for obtaining the privacy protection assurance provided by an entity consists of evaluating the privacy requirements fulfilled by the entity and determining its associated assurance category. A general, hihg-level framework is proposed in Table 5.19.

Privacy Assurance (LOP)	**Description**
High	If a range of 70% 100% of optional privacy requirements is fulfilled apart from the mandatory ones
Medium	If a range of 35% 70% of optional privacy requirements is fulfilled apart from the mandatory ones
Low	If the set of minimum privacy requirements is fulfilled

Table 5.19: Privacy assurance framework

As it can be seen, the framework definition provided in this case abstracts the underlying criteria. The reason is that, there is ongoing research work on defining a list of privacy requirements to be used for ranking levels of privacy assurance. More specifically, in the field of FIM, Kantara recently started a working group [9] to develop a privacy framework.Their idea is to to provide a description of the privacy components and derive a set of levels of privacy that can be widely accepted in the same way as LOAs. To give an idea of the kind of criteria to be evaluated for the privacy metric, we list here some of these criteria according to the documentation publicly available from Kantara:

- Purpose of processing personal data
- Relevance of attributes
- User consent
- Informed consent
- Data protection directives
- Release of data when the entities operate in different countries

[9] http://kantarainitiative.org/confluence/display/p3wg/

3. What is the qualitative scale?

As in the case of the metrics defined so far, we adopt a three level qualitative scale that encompasses *Low*, *Medium*, and *High* levels. Each of these levels represents the assurance achieved in privacy depending on the strength of the mechanisms used for this purpose.

- **Pre-Fed->Interoperability->Technical**
- **Pre-Fed->Interoperability->Operational**
- **Pre-Fed->Interoperability->Legal**

We group the definition of the interoperability metrics, since the frameworks are very similar.

1. What does the metrics measure?

In this categories we define the $\mathrm{INTEROP_T}$, $\mathrm{INTEROP_O}$ and $\mathrm{INTEROP_L}$ metrics, which measure the degree of technical, operational and legal interoperability between the involved parties, respectively.

2. How is the measure performed?

The procedure for obtaining the interoperability metrics consists of evaluating the technical, operational and legal requirements fulfilled by the entity and determining its associated assurance category. We define general frameworks for this purpose in the same way as the frameworks for LOP and ACC were defined (see Tables 5.18 and 5.19). That is, assurance categories are obtained based on the coverage of mandated and optional requirements.

Since interoperability means that two entities are able to work together from a technical, operational or legal point of view, the lower level implies as a minimum that interoperability is guaranteed. Higher levels in the categorization contemplate additional aspects where entities are interoperable.

Regarding technical aspects, SAML leaves deployers a lot of options, like how to pass attributes, what information should be signed/encrypted or what binding to use. An unwanted effect of all these available choices is that two SAML deployments may not work as smooth together as expected. As an example of minimum set of bindings and rules that needs to be followed we can cite [Solberg, 2011], which specifies behavior and options that deployments of the SAML V2.0 Web Browser SSO Profile are required or permitted to rely on. For the case of operational interoperability, a lot of options are possible as well, such as metadata caching intervals, frequency of certificate renewal, procedures for certificate validation, etc. Finally, for the case of legal interoperability, there are also a number of options. For example, special directives may be required when dealing with health related data. Thus, each entity using these frameworks should define its minimum set of requirements at each interoperability dimension.

In the case of a SAML-based system, metadata, SLAs and policies may be used as source data for the interoperability metrics if appropriate mechanisms are defined. A full list of possible interoperability requirements should be investigated.

3. What is the qualitative scale?

As in the case of the metrics defined so far, we adopt a three level qualitative scale that encompasses *Low*, *Medium*, and *High* levels. Each of these levels represents the assurance achieved in technical, operational or legal interoperability assurance depending on the fulfilled requirements.

- **Pre-Fed->Knowledge->Direct**

1. What does the metric measure?

In this category we define the $\mathrm{KNOW_D}$ metric, which measures the direct knowledge assurance about the other party.

2. How is the measure performed?

The procedure to obtain the direct knowledge assurance is simpler than for the rest of the metrics, and it consists of checking the internal data to determine if the evaluating entity has an entry for the entity under evaluation in its database or trust lists. There are only two possible outcomes for this process and thus, the metric is binary.

The source for this metric are the internal data structures (e.g., trust lists) where information about external entities is stored.

3. What is the qualitative scale? In this case we use a binary scale with values *True* and *False* for the cases where direct knowledge exists or no, respectively.

- **Pre-Fed->Knowledge->Indirect**

  1. What does the metric measure?

  In this category we define the $\mathrm{KNOW_I}$ metric, which measures the indirect knowledge about the other party.

  2. How is the measure performed?

  The procedure to obtain the indirect knowledge assurance is the the same as for the direct knowledge metric. In this case, it consists of checking the internal data to determine if the evaluating entity has indirect information about the entity under evaluation in its database or trust lists. There are only two possible outcomes for this process and thus, the metric is also binary.

  The source for this metric are the internal data structures (e.g., trust lists) where information about external entities is stored.

  3. What is the qualitative scale?

  In this case we use a binary scale with values *True* and *False* for the cases where indirect knowledge exists or no, respectively.

- **Post-Fed->Service Specific**

  This category encompasses metrics related to the specific service in which the current transaction is performed. The metrics in this category are related to Service Level Objectives (SLOs), i.e., metrics which define characteristic of a service in precise, measurable terms. The LOA is a well-known SLO in identity services. Here we include only this metric under the service specific category, but each service should define its own specific metrics (e.g., throughput, bandwidth, etc.)

  Since the LOA metric was already described at the beginning of the section, the description is not included here.

- **Post-Fed->Historical Interactions**

  1. What does the metric measure?

Measures the degree of confidence that the collaborating entity will operate as expected in the context of the current transaction.

2. How is the measure performed?

The procedure for obtaining the metric consists of: 1) calculating the *a posteriori* probability of a satisfactory interaction, and 2) assimilating it to an assurance value.

For the first step, based on the number of satisfactory (*sat(C_i)*) and unsatisfactory interactions (*unsat(C_i)*) with the entity under evaluation, the *a posteriori* probability of a satisfactory interaction ($p^{(+)}(C_i)$ is obtained by applying formula 5.1

$$p^{(+)}(C_i) = \frac{sat(C_i)}{unsat(C_i) + sat(C_i)} \tag{5.1}$$

Where C_i refers to the type of the current transaction, assuming an entity can operate in a set of transaction types $C = C_1,\ C_2,\ \ldots,\ C_n$.

Based on this, the procedure for obtaining the HINT metric consists of determining its associated assurance category. Regarding the kind of criteria used to categorize each level of assurance, we propose a simple framework, which is depicted in Table 5.20.

We define a general framework for this purpose in Table 5.20.

Historical Interactions (HINT) Assurance	**Description**
High	If $p^{(+)}(C_i) \geq 99\%$
Medium	If $97\% \geq p^{(+)}(C_i) < 99\%$
Low	If $95\% \geq p^{(+)}(C_i) < 97\%$

Table 5.20: Historical Interactions assurance framework

In order to have a source data to obtain the HINT metric, it is required that entities implement a monitoring system to assess transactions and count the number of satisfactory and unsatisfactory ones.

3. What is the qualitative scale?

As in the case of the metrics defined so far, we adopt a three level qualitative scale that encompasses *Low*, *Medium*, and *High* levels. Each of these levels represents the assurance related to the history of interactions depending on the probability that the current transaction is satisfactory.

After completing the risk metric derivation procedure based on the taxonomy, the set of basic metrics that we have defined are summarized in Table 5.21:

Category (L2)	Metric Name	Description
Security and Privacy	$\mathrm{CONF}_{\mathrm{TL}}$, $\mathrm{CONF}_{\mathrm{ML}}$	Measure confidentiality of information exchanged at transport level and message level, respectively.
	$\mathrm{INT}_{\mathrm{TL}}$, $\mathrm{INT}_{\mathrm{ML}}$	Measure integrity of information exchanged at transport level and message level, respectively
	$\mathrm{AUTH}_{\mathrm{TL}}$, $\mathrm{AUTH}_{\mathrm{ML}}$	Measure authentication of information exchanged at transport level and message level, respectively.
	NON-REP	Measures the degree of Non-repudiation.
	AV	Measures the degree of availability.
	ACC	Measures the degree of accountability.
	LOP	Level of Protection measures the degree of data protection that either an IdP or a SP provides for identity information entrusted to them by a user.
Interoperability	$\mathrm{INTEROPT}_{\mathrm{T}}$	Measures the degree of technical interoperability between the involved parties
	$\mathrm{INTEROP}_{\mathrm{O}}$	Measures the degree of interoperability between the operational policies of the involved parties
	$\mathrm{INTEROP}_{\mathrm{L}}$	Measures the degree of interoperability between the legal policies of the involved parties
Knowledge	$\mathrm{KNOW}_{\mathrm{D}}$	Measures the degree of direct knowledge about the other party
	$\mathrm{KNOW}_{\mathrm{I}}$	Measures the degree of indirect knowledge about the other party
Service Specific	LOA	Level of Authentication Assurance (LOA) measures the degree of confidence in identifying an entity to whom a credential was issued
Historical interactions	HINT	Measures the degree of confidence that the collaborating entity will operate as expected in the context of the current transaction

Table 5.21: Basic set of metrics for risk quantification in FIM

Table 5.21 shows the set of basic metrics derived from the proposed taxonomy and their semantic definition. From this basic set, the aggregated metrics will be developed in the next section. Furthermore, a comprehensive catalogue with the definition of all the metrics can be found in Appendix B. So far, we have provided a semantic high-level definition of the metrics, and assigned a linguistic scale to each one. In the following, we will explain how to map this qualitative scales into numeric values and aggregate them to obtain final risk figures upon which decision making is possible.

5.4. Risk Assessment

The purpose of this section is to present the aggregation problem in more detail and complete the whole quantitative risk model for identity federation. Thus, after the identification of risks terms and analysis in sections 5.2 and 5.3, we proceed now with the evaluation or aggregation. As stated before, the main goal is to obtain a representative value that can be used in decision making. And this value depends on multiple criteria that we have identified step by step with the design of the taxonomy. Considering these features, the aggregation problem perfectly fits in the framework of multicriteria decision making (MCDM) [Triantaphyllou, 2000].

The MCDM theory is a sub-discipline of operations research that explicitly considers multiple criteria in decision-making environments based on the premise that structuring complex problems well and considering multiple criteria explicitly leads to more informed and better decisions. In summary, MCDM approaches are a tool to integrate multidimensional evaluations and frequently rely on decision hierarchies (such as our taxonomy), features that fit with our goals and design and which led us to build our risk evaluation model under this theory.

There are two main approaches of MCDM, namely multiattribute utility theory (MAUT) [Keeney and Raiffa, 1993], and the preference modeling approach [Fodor and Roubens, 1994]. In multiattribute utility theory, an absolute score is given to each alternative with respect to each criterion, and the global score, taking into account all the criteria, is obtained by aggregating all the partial scores. By contrast, in preference modeling, a preference degree is assigned to every ordered pair of alternatives, with respect to each criterion. Then, a global preference degree is obtained by aggregating all the partial preference degrees. Due to the nature and different semantics of the quantities to be aggregated in MAUT and preference modeling (i.e., scores or preference degrees), the approaches are also referred as “cardinal approach” and “relational approach”, respectively.

Another interesting approach comparable to preference modeling is the usage of Multi Dimensional Scaling (MDS) techniques [Borg and Groenen, 2005]. MDS allow to rank entities based on the calculation of their similarities or dissimilarities regarding different dimensions. In fact, there is work that builds on MDS to solve, e.g., the problem of selecting the most appropriate network or peer to interact with based on a set of criteria (cost, distance, security) [Díaz Sánchez, 2008] [Díaz et al., 2006].

In our case, the decisions to be made in a federation scenario are, as previously stated in the chapter goals, the following:

- whether to collaborate or not with another entity
- which entity in a set is the best alternative for cooperation

For the first kind of decision, absolute evaluation is required since we do not have alternative options for comparison. Consequently, the MAUT technique is more suitable to aggregate the multiple risk dimensions and to obtain a final value upon which to decide. In the second kind of decision, both preference degrees and absolute scores may be used, so any approach is applicable. For these reasons, we adopt the MAUT approach and adapt it to our application, since it allows us to cover both kinds of decisions.

Whatever the MCDM approach to be taken, the involved phases are common [Marichal, 1998]. Thus, multicriteria decision making procedures consist of three main steps:

1. **Modeling phase:** In this phase we look for appropriate models for constructing the partial scores/preference degrees and also for determining the importance of each criterion (i.e., the weights).
2. **Aggregation phase:** In this step we try to find a unified (global) score for each alternative, on the basis of the partial scores and the weights.
3. **Exploitation phase:** In this phase we transform the global information about the alternatives either into a complete ranking or into a global choice of the best alternatives.

To describe MAUT formally we adopt a terminology similar to that in [Marichal, 1998]:

- $P = A,\ B,\ C,\ .\ .\ .$ is a non-empty set of objects or alternatives (in our case the providers), among which the decision maker must choose.
- We have a collection of criteria $N = 1,\ .\ .\ .\ ,\ n$ we desire to satisfy. Each criterion i is represented by a mapping g_i from the set of alternatives P to a measurement scale $S_i \in \Re$. The value $g_i(A)$ is then called the partial score of alternative (provider) A with respect to criterion i.

- The value *$g_i(A)$* is then called the partial score of A with respect to criterion i. In most applications, it is assumed that each S_i is the unit interval [0, 1].

According to this notation, the global score of an alternative A using MAUT is calculated by means of an aggregation operator which takes into account the weights associated to criteria. Although a wide diversity of techniques have been recommended for resolving multiattribute decision problems, most are ultimately based on a weighted linear model. Thus, the linear aggregation is used in MAUT as shown in expression (5.2):

$$\sum_{i}^{n} \omega_i \cdot g_i(A) \tag{5.2}$$

Here, *$g_i(A)$* is the evaluation of object A on the *i-th* value dimension, ω_i is the weight determining the impact of the *i-th* value dimension on the overall evaluation (also called the relative importance of a dimension), and n is the number of different value dimensions.

The key points of the method are: identifying what is important for the evaluation (dimension hierarchy), identifying how well each alternative does on each criterion (score g_i), and identifying the importance (weights ω_i). In the following, we explain how we fulfill these points and how we apply and adapt the MAUT method to construct the whole risk model.

Firstly, the taxonomy previously developed in Section 5.2 contributes to the modeling phase in MAUT by providing the criteria to be assessed. In our model, the criteria are the different assurance dimensions contemplated in the taxonomy. There are two contexts to assess risk, namely Pre-Federation and Post-Federation. The criteria and sub-criteria to be taken into account in each case are modeled by the hierarchy under them. Having identified the criteria, we define the mapping functions $\mathrm{g_i}$ based on the assurance metrics. For this purpose, we establish a qualitative to quantitative correspondence for the defined assurance scales of our metrics according to Tables 5.22 and 5.23:

Qualitative Assurance Value	**Quantitative Assurance Value**
High Assurance	3
Medium Assurance	2
Low Assurance	1
No Assurance	0

Table 5.22: Quantitative mapping of qualitative assurance scale levels for ordinal metrics

Qualitative Assurance Value	**Quantitative Assurance Value**
True	1
False	0

Table 5.23: Quantitative mapping of qualitative assurance scale levels for binary metrics

As shown in Table 5.22, numbers have no other meaning that defining an order relation on the scale. In Table 5.23, however, the quantitative numbers provide binary semantics indicating if the criteria is satisfied or not. Since there are different scales involved in the problem, normalization is required to conduct the ulterior aggregation. Consequently, we define the mapping function $\mathrm{g_i}$(A) to obtain the normalized scores for each criteria as in formula 5.3:

$$g_i(A) = \frac{AssuranceValue - Min(AssuranceValue)}{Max(AssuranceValue) - Min(AssuranceValue)} \tag{5.3}$$

As we can observe, the values move now on the interval [0,1].

Accordingly, we give a couple of examples considering the criteria at taxonomic level L2:

- If a provider A in the set of alternatives P has a high privacy assurance (LOP = High), its corresponding partial score for this criterion would be $g_i(A)$= $^{3\text{-}0}/_{3\text{-}0}$ = 1.
- In case the decision maker has direct knowledge about a provider A (KNOW$_D$ = True), its corresponding score for this criterion would be $g_i(A)$ = $^{1\text{-}0}//1\text{-}0$ = 1.

Formally, each provider $A \in P$ can be assimilated with the vectors of its partial scores (i.e., [$\mathrm{g_1}$(A), . . . , $\mathrm{g_n}$ (A)] $\in \mathrm{S_1} \times \ldots \times \mathrm{Sn_n}$.) For example, according to the security dimensions, a provider A with a vector [$^3/_3$ 0 0 $^2/_3$ $^1/_3$ $^3/_3$ $^3/_3$] means that its assurances in confidentiality, integrity, authentication, non-repudiation, availability, accountability, privacy are 3, 0, 0, 2, 1, 3 and 3, respectively. The scores are then obtained depending on the features of the providers based on the metric framework described in the previous section.

Having defined the scores, the next step to finish the modeling phase and start with the aggregation, consists of determining the weights of each criteria. Since the criteria may not have the same importance, it is then useful to define a weight ω_i associated to each criterion i. Such a weight represents the strength or importance of this criterion. We aim to detail the election of the weights by doing aggregation tests applying the MAUT expression in 5.2. In this regard, it is to note that our hierarchies (see figure 5.1) have multiple levels, i.e., there are criteria and also sub-criteria under them. However, in contrast with other mechanisms, MAUT does not support sub-criteria hierarchies directly. But it can instead be applied recursively until obtaining a final value. Consequently, due to the multi-level nature of the taxonomy, this is the approach followed here. Accordingly, the adapted formula we use for aggregation is:

$$Agg^{k,j}(A) = \sum_{i}^{n} \omega_i^{k+1,j} \cdot g_i^{k+1,j}(A) \tag{5.4}$$

Which indicates that the aggregated assurance value for a provider A with respect to criterion j at level k (Agg_j^k(A)) is the weighed summation of the assurance values of A for all the criteria i at level k+1 that are a sub-criterion of j. The weights $\omega_i^{k+1,j}$ represent the relative importance of the criteria and hold $\sum_i^n \omega_i^{k+1,j} = 1$

Using vectorial notation we denote:

- $\mathrm{WV}^{k,j} = [\omega^{k,j}{}_1, \ldots, \omega^{k,j}{}_n]$ as the *Weights Vector* at level k regarding criteria j.
- $\mathrm{SV}^{k,j}(\mathrm{A}) = [\mathrm{g}^{k,j}{}_1(\mathrm{A}), \ldots, \mathrm{g}^{k,j}{}_n(\mathrm{A})]$ as the *Score Vector* for provider A at level k regarding criteria j.

So expression 5.4 can be rewritten as in 5.5:

$$Agg^{k,j}(A) = WV^{k,j} \times SV^{k,j}(A)^T \tag{5.5}$$

The methodology then consists of recursively reducing the problem by aggregating sub-criteria into global values until we have a single unique value or criterion. Thus, the hierarchy is used in a recursive bottom-up way. The intermediate aggregated values are called aggregated assurance metrics and these values are the inputs ($g_i(A)$) used for aggregation at the immediate upper level. Furthermore, we decided to apply an additional treatment at each iteration to include risk policies and to satisfy minimum requirements. The adaptation of the method and election of weights are shown by example in the validation chapter (see Chapter 8).

5.5. Conclusions

In this chapter we have defined the risk assessment model that is implemented by the *Risk Manager* component of the proposed architecture for dynamic federation. The model is based on the MAUT theory, which allows to combine criteria of different nature into a single value to be used in decision-making. The criteria in this case are the dimensions of risk. Since the process of identifying the dimensions of risk in FIM is not a trivial task, we started by designing a taxonomy to capture the different aspects of a FIM relationship that may contribute to risk. Based on this analysis, we derived the set of quantitative metrics to be aggregated following the MAUT approach.

In regard to aggregation, the risk dimensions are weighed according to the preferences and risk policies of the evaluating entity. The adjustment of the weights is presented in Chapter 8 together with the validation tests. Finally, a catalogue containing detailed information about all the proposed risk metrics can be found in Appendix B.

Chapter 6

Trust and Reputation Model proposal

"Without trust we cannot stand."

Confucius

Contents

6.1. Chapter Overview

The difficulty of capturing the notion of trust in formal models in a meaningful way has sometimes led to reject it as a computational concept. However, the formal definition of trust is required. We need to formalize it because our lives are so technically-mediated that we need devices and applications to act as our proxies, and to act on the basis of the same concepts we ourselves rely on in our daily lives. Having a formal model to compute and represent trust as a number provides a basis for easy implementation and automation. With these premises as foundation, this chapter focuses on developing the trust model that is to be implemented as part of the architecture for dynamic federation. Section 6.2 defines the trust metric, detailing the evidences used, and how they are combined to obtain a quantitative value. Basically, authentication information is merged with behavior data, i.e., reputation or history of interactions. Next, Section 6.2 elaborates on the mechanisms defined to share reputation data in the FIM ecosystem. Finally, the main conclusions of the chapter are summarized in Section 6.4.

6.2. Designing a Trustworthiness Metric for Trust Management

Since their origins trust management systems have been used in order to assist entities that have to interact with others in a system, being a useful tool for the decision-making process. In order to establish the trust relationship a trust management system is usually composed of a symbolic language for representing trust and a way of measuring trust (trust metrics), that derives the trust assessment. In this sense, the computational formalization of trust plays a crucial role. Having a formal model to compute and represent trust as a number provides a basis for easy implementation and automation, as well as a common understanding of what is measured. As pointed by Marsh in his widely cited PhD thesis [Marsh, 1994], the expected benefits of a computational trust formalism are increasing reliability and performance of electronic communities, and an achievement of more cooperation in open and less protected environments.

In the particular case of FIM systems nowadays, trust management models basically consist on the pre-configuration of lists in such a way that if an entity is contained on (or reachable through) the list, then it is trusted. These mechanisms are thus static and highly dependent on administrative configuration tasks. And since decisions are the result of binary assessments, the flexibility is limited.

Furthermore, there are no formal computational models that map these FIM trust procedures to numerical expressions. Here we aim to fill this gap by formalizing a model that captures the features of current FIM systems and introduces new dimensions to add flexibility and richness. This enhanced model will favor better informed decision-making and foster collaboration with previously unknown entities. The formalization of a trust model in a specific domain is a multi-stage process. For the definition of the trust management model, we will follow these steps:

1. **Trust evidences identification:** Since the design of any trust management solution is highly influenced by the problem being addressed, the first step consists in identifying and selecting the proper input data, i.e., the trust evidences in the FIM context. We will conduct an analysis to identify the existing evidences that are being used, but also introduce new ones.
2. **Quantitative mapping:** The trust evidences are represented in a quantitative scale, whose semantics are also detailed. For this purpose, we rely on the terminology and formalizations in the literature and build on them.
3. **Evidences combination:** A trust computation is performed over the evidences to produce a trust value, which reflects the estimation of the trustworthiness of an entity.

The result after the completion of the above steps is a trustworthiness metric, which will be the basis for making trust decisions. The explanation of its usage within the decision making procedures completes the definition of the whole model.

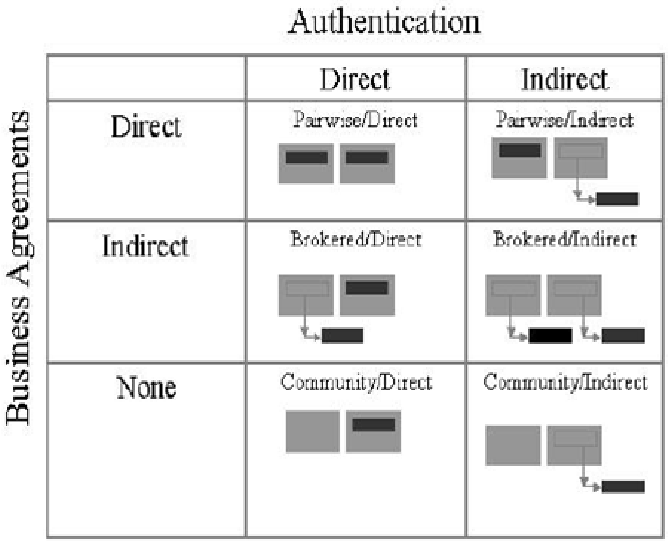
6.2.1. Trust Evidences Identification

The most complete document that conceptually describes and provides guidelines for trust models in the FIM context is [Boeyen et al., 2004]. Thus, we base on this document to conduct the analysis of the initial trust evidence space. The document describes different alternatives for trust establishment based on the notions of business trust and authentication trust. The following definitions, contained in the document, are the basis to understand the trust models:

- **Trust Anchor List (TAL).** Entities accepting cryptographic authentication of other entities will maintain trust anchor lists, identifying the entities and associated keys (typically within digital certificates) that they trust for authentication purposes and upon which validations will be based. If indirect authentication is accepted, the TAL must contain the intermediary entities through which an authentication path can be derived. Entities in a TAL are called Trust Anchors.

- **Business Anchor List (BAL).** Entities requiring business agreements in order to interoperate with other entities will maintain business anchor lists identifying the entities with which direct business trust relationships have been established. If an entity accepts indirect business agreements, its BAL must contain the intermediary entities through which a business agreement path can be derived. Entities in a BAL are called Business Anchors.

Based on the above definitions, the possible taxonomy of contemplated cooperation models is shown in Figure 6.1



	Direct	Indirect
Direct	Pairwise/Direct	Pairwise/Indirect
Indirect	Brokered/Direct	Brokered/Indirect
None	Community/Direct	Community/Indirect

Figure 6.1: Trust model taxonomy for SAML-based FIM systems (©[Boeyen et al., 2004])

On the one hand, regarding business, entities may operate under direct agreements, indirect agreements, or without any agreement. On the other hand, direct or indirect authentication should exist. The combination of these possibilities leads to pairwise, brokered or community models, with direct or indirect authentication.

Accordingly, the process to determine whether a requesting entity can be trusted or not, could be summarized with the work-flow presented in Figure 6.2.

For an entity A to determine whether a suitable basis exists to carry out trusted transactions with another entity B, it operates on the following data: B's identity, A's TAL, A's BAL, and A's operational policies, indicating the types of paths it accepts. And the

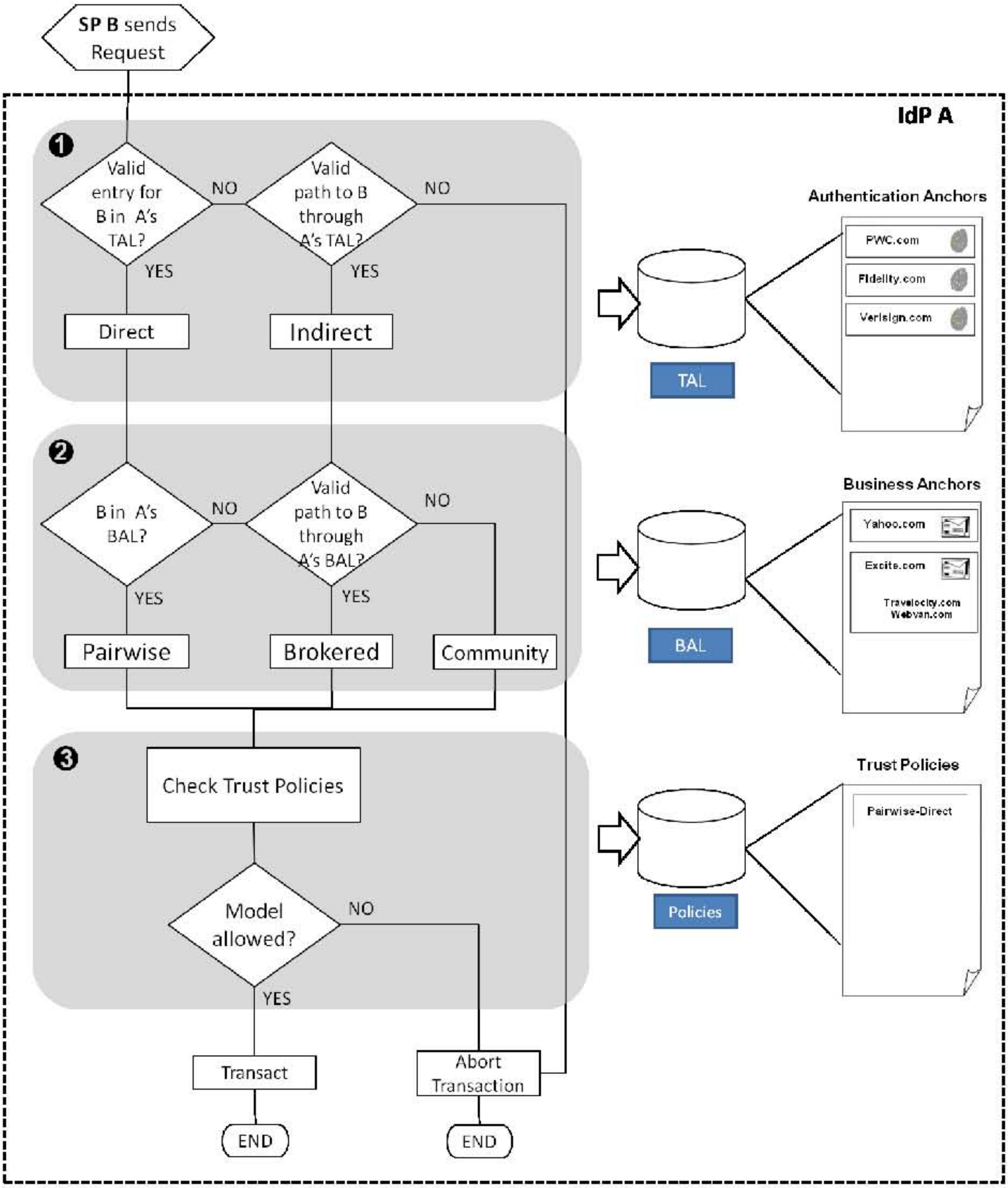


Figure 6.2: Trust Management model work-flow (©[Boeyen et al., 2004])

necessary processing at a conceptual level, which starts when A receives a transaction request from B, consists of the following steps:

1. **Validation of an authentication trust path.** This process begins by determining whether A's TAL contains an entry for B. If so (e.g., in the Figure 6.2 example, if B's identity is Fidelity.com), Direct Trust applies, and A possesses the key required to authenticate messages and/or connections received from B. If not, A must determine whether one or more of the entries in its TAL enables it to construct an authentication path to B. If an authentication path can be constructed and validated, Indirect Trust applies, and A can traverse that path to obtain the key required to authenticate messages and/or connections received from B. If no path can be constructed, then A is unable to authenticate B.

2. **Validation of a business agreement path.** This process begins by determining whether A's BAL contains an entry for B. If so (e.g., in the Figure 6.2 example, if B is Yahoo.com), Pairwise Trust applies. If not, A must determine whether one or more of the entries in its BAL enables it to construct a business agreement path to B. If a business agreement path can be constructed (e.g., in the Figure 4 example, a path to Travelocity.com via Excite.com), Brokered Trust applies. If not, no business agreement applies between A and B, and any transactions must be carried out based on a Community Trust model.

3. **Policy checking.** At this stage in the process, A has identified the "shortest" applicable type of authentication path (Direct or Indirect) and of business agreement path (Pairwise, Brokered, or Community) reaching to B. It must now determine whether these paths satisfy its policies and, if so, whether they dictate any limits or constraints on the transactions that it will be willing to undertake with B; a peer reachable via Pairwise Trust, e.g., might be accorded broader rights than one reachable only at the Community Trust level.

There are two important constraints in the procedure that make it rigid and static. Firstly, if a key/certificate or a path to a key/certificate is not available through the TAL, then transaction is directly aborted. There is no other mean to obtain the key/certificate dynamically and assign a trust level to it. Secondly, it is assumed that the addition and removal of entities to both TAL and BAL lists are serious decisions that should normally happen only as a result of explicit administrative actions.

Consequently, there are two sources for deriving trust, namely TAL and BAL. Also, it can be noted that the BAL is always a subset of the TAL. That is, an entity only conducts transactions with direct or indirect authenticated entities, but business contracts may exist or not. Since being in the TAL is indispensable for trusting an entity, it constitutes, from our point of view, the main trust evidence or input to compute the trustworthiness metric. The agreements on the BAL however are related to the risk computation and they are already contemplated on the risk part of our architecture.

But the authentication trust alone is limited. Having installed a valid certificate means that we have assurance that the entity willing to cooperate is the holder of the private key associated to the public key contained in the certificate. We know this binding is authentic; the entity is who it claims to be and we have a cryptographic tool to establish secure communications. However, we know nothing about its actual behavior. For this reason, we argue that other trust evidences containing this behavior information are desirable to complement our knowledge. Behavior can be inferred from previous transactions, if they exist; or from reputation, in case no previous history of direct interactions is available.

Reputation data has proven useful as a mean to convey empirical information and improve trust, providing a notion of behavior when no direct knowledge exists. Numerous studies in the field of distributed computing demonstrate this fact. However, the application of this dimension of trust to FIM scenarios has not been yet fully addressed, and here we aim to show and evaluate its utility.

Summarizing, the trust dimensions we consider to be the basis of the FIM trust model are: **Authentication Trust** (T_{Auth}) and **Behavior Trust** ($\boldsymbol{T_{Behav}}$) , as depicted in Figure 6.3. The trust evidences taken as input data at each dimension are the digital certificates and the transaction history or reputation, respectively.

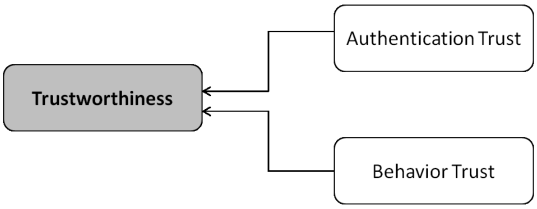


Figure 6.3: Dimensions of the trustworthiness metric

6.2.2. Quantitative Mapping and Evidence Combination

Computational trust models provide accurate trust assessment based on the usage of numeric values. Here we map the conceptual procedure for trust management in current FIM described in Figure 6.2 to a numeric computational procedure[1]. Next, over this starting point, we include extensions to make the trust model more flexible.

The common implementation of TALs is PKI-based [Adams, C. and Farrell, S., 1999], i.e, the TAL contains digital certificates and certificate validation procedures are performed over the list. From the trustworthiness point of view, PKIs are binary systems: a certificate is either trustworthy or not trustworthy. Certificate validation (thoroughly detailed in [Freeman, T. and Housley, R. and Malpani, A. and Cooper, D. and Polk, W., 2007]) consists of checking its integrity, expiration status and revocation status. If these aspects are valid and the issuer is considered a trusted source, then the certificate is trustworthy. If the validation fails at any of these aspects, the certificate is not trustworthy. For path validation, all the certificates in the path must be valid. On the other hand, commonly used values in the trust literature for full trust and full distrust are 1 and 0 respectively. Considering this, the quantitative mapping to computationally formalize the "validation of an authentication path" procedure is:

- If a direct entry exists in the TAL with a certificate for B ant the certificate is valid, then authentication trust is equal to 1. If the certificate is not valid, then authentication trust is equal to 0.
- If an indirect entry exists in the TAL that allows the construction of an authentication path to B, and the whole path is valid, then authentication trust is equal to 1. If the certificate path is not valid, then authentication trust is equal to 0.
- If no direct certificate entry or certificate path can be constructed through the information in the TAL, then authentication trust is 0.

As it can be observed, the metric is purely binary. Furthermore, following this procedure, unknown entities are considered untrustworthy and they are not given a chance to operate. In a similar way, pre-configured entities are considered fully trusted without knowing anything about their actual behavior. This assignment of numeric values to T_{Auth} is

[1] Only the process for obtaining authentication trust is mapped, since the business trust will not be used in the trust model but in the risk one

depicted in Figure 6.4.

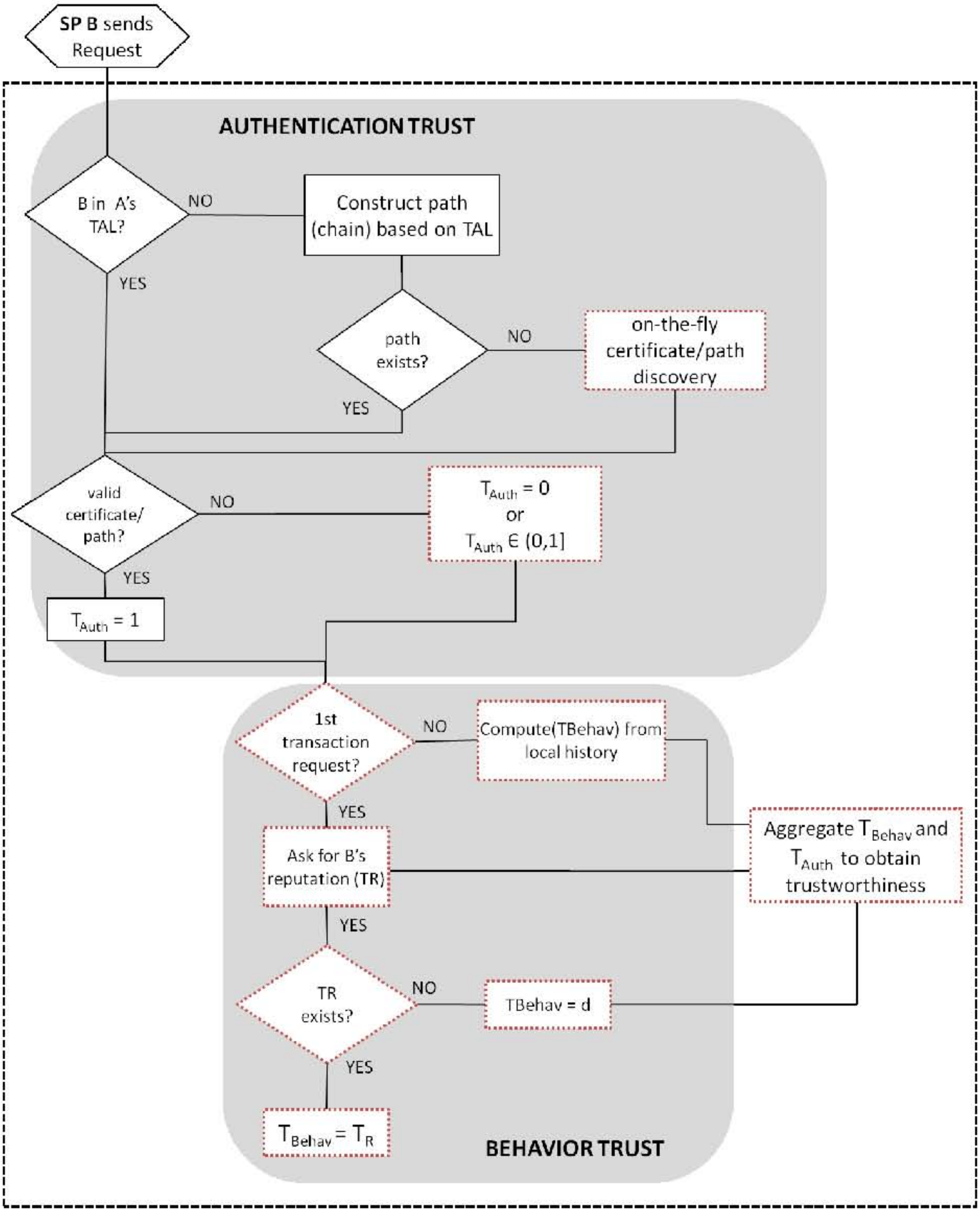

SP B sends Request
AUTHENTICATION TRUST
B in A's TAL?
NO
Construct path (chain) based on TAL
YES
path exists?
NO
on-the-fly certificate/path discovery
YES
valid certificate/ path?
NO
$T_{Auth} = 0$ or $T_{Auth} \in (0,1]$
YES
$T_{Auth} = 1$
1st transaction request?
NO
Compute(TBehav) from local history
YES
Ask for B's reputation (TR)
Aggregate T_{Behav} and T_{Auth} to obtain trustworthiness
YES
TR exists?
NO
TBehav = d
YES
$T_{Behav} = T_R$
BEHAVIOR TRUST

Figure 6.4: Proposed work-flow and quantitative mapping for the FIM trust management model

But to get a better measure of trustworthiness a continuous scale must be used that allows precise representation of computed trust values. We propose to use trust values in the range [0, 1], and -1 to denote lack of information. We propose to improve the trust metric calculation procedure in three ways:

1. Allowing non-binary values for T_{Auth}
2. Allowing the dynamic inclusion of certificates with no trusted anchor in the TAL
3. Adding behavior information

These three aspects are depicted with dotted boxes in the flow diagram of Figure 6.4.

Firstly, to allow more flexibility and granularity in the validation, alternative assessment mechanisms may be used. For example, if the status of the certificate cannot be checked due to the temporal unavailability of the revocation server, the validation value is assigned a value proportional to the probability that it is revoked instead of being considered not valid. This technique is called probabilistic validation. Thus, we only have a 0 value when the certificate is really not valid, a 1 value when it is totally valid, and numeric values in between when the validity can be assured to a certain level. Proposals that map this certificate validity assurance to quantitative numbers exist in the literature [Haenni, R., 2005]. It can be noted that, by applying this approach, T_{Auth} will have a value in the continuous range [0,1].

Secondly, the models in Figure 6.1 assume that an entity that is unreachable from the TAL is untrustworthy. However, as remarked in [Boeyen et al., 2004], entities may be able to establish trust between them through exchange of trust metadata. Thus, our model includes a step to directly ask for the exchange of certificates on the fly when no Trust Anchor on the TAL is useful.

After T_{Auth} is computed according to the above procedures, the third step consists of gathering behavior information in order to compute T_{Behav}. Next, both values are combined into a trustworthiness metric to be used in decision-making. At this point, the assignment of quantitative values to T_{Behav} is performed this way:

- If there is no previous interaction between the parties, i.e., the current transaction is the first, then the evaluating entity asks for reputation (T_R) about the requesting entity. If a reputation value is obtained, then $T_{Behav} = T_R$. In the extreme case that T_R is equal to -1, i.e., no reputation information exists, T_{Behav} is assigned a trust disposition value d. This value, in the range [0,1], indicates the disposition of the entity to trust another one when there is no data about its behavior.

- If previous interaction between the parties exists, i.e., the current transaction is not the first, the value of T_{Behav} based on the history of transactions until the current one is used. For this purpose, after each transaction, the behavior value must be updated taking into account the existing trustworthiness value and the satisfaction on the current transaction (*sat(i)*). Several models have been proposed in the literature to model the evolution of trust as a function of the satisfaction on the transactions. We recommend and base on the mathematical trust evolution scheme used in the PTM [Mendoza et al., 2011] model.

After T_{Behav} is calculated, the global trustworthiness is computed as the product $T_{Auth} x T_{Behav}$. Mathematically, the expressions to model these procedures are shown in equations 6.1 and 6.2, based on the trust parameters we have defined (summarized in Table 6.1).

$$Trustworthiness(i) = T_{Auth}(i) \cdot T_{Behav}(i-1) \tag{6.1}$$

$$T_{Behav}(i) = \begin{cases} \begin{cases} T_R \;\; if \;\; T_R \neq -1 \\ \;\; d \;\; otherwise \end{cases} \; when \quad i = 0 \\ f_{evol}(trustworthiness(i), sat(i)) \;\; when \;\; i > 0 \end{cases} \tag{6.2}$$

Parameter	Description
trustworthiness(i)	Measures the trust that can be placed on an entity after i transactions
$T_{Auth}(i)$	Authentication Trust at transaction i
$T_{Behav}(i)$	Behavior Trust after transaction i
$T_R(i)$	Reputation Trust
d	Trust disposition
f_{evol}	Trust evolution function
sat(i)	Satisfaction on transaction i

Table 6.1: Parameters for the FIM trust management model

Trust values in Table 6.1 (i.e., *trustworthiness*, T_{Auth}, T_{Behav}, T_R) take values in the continuous scale [0,1]. The trust disposition factor *d*, also in the scale [0,1], as well as the function f_{evol} for the evolutionary model, are selected by the evaluating entity and configured in its policies. The value range for *sat(i)* will depend on the f_{evol} used in order to update T_{Behav} after each transaction. Typically, discrete values 0 and 1 are used to express satisfying and unsatisfying transactions, respectively.

Now that the process to compute trust has been defined, we elaborate on the mechanisms to obtain and aggregate the reputation data in a FIM scenario.

6.3. Handling Reputation

The main novelty of the trust management model proposed here is the inclusion of reputation information. Reputation is central to all kinds of human interaction, including interpersonal relationships, international diplomacy, stock markets, etc. Computationally, reputation has been introduced in Internet applications proving useful to improve collaboration in environments where uncertainty is present. Specially, in P2P networks (e.g., for file sharing) there has been an intensive research in the last decade [Hoffman et al., 2009]. Here we aim to adapt the existing knowledge on decentralized reputation protocols to adapt and define a solution that is applicable in FIM scenarios. Far from reinventing the wheel, we first analyze the best know reputation protocols, and then select the more suitable and adapt it.

A reputation protocol is generally composed of a component for gathering behavioral information, and a component for scoring entities. In turn, each component requires a combination of mechanisms to function. For defining the protocol we follow the recommendations in the RFC "*Writing Protocol Models*" [Rescorla, 2005]. Accordingly, a protocol model is described by answering three basic questions: 1.) What problem is the protocol trying to achieve?, 2.) What messages are being transmitted and what do they mean?; and 3.) What are the important, but unobvious, features of the protocol?. The response to the first question, already motivated, is to include reputation data in FIM environments with the goal of fostering more dynamic and secure collaborations. Next, we develop the protocol model answering questions 2 and 3 in the following sections. During the process we also highlight the differences with current protocols, and what parts are new, specific and necessary in the context of FIM.

6.3.1. Protocol Overview

As a previous knowledge base for defining the reputation protocol in the FIM context, we first reviewed the most notable systems. Perhaps the most widely used reputation system is that of eBay, which consists of a single trusted entity that collects all transaction reports and rates each user. But since it is a centralized system, it is not applicable to our scenario. In the envisioned FIM scenarios, there is not a single trusted third party to collect the ratings, but entities belonging to multiple unknown domains willing to cooperate. Thus, decentralized reputation systems, as the proposed for P2P environments, are more suitable.

Regarding P2P reputation systems, two proposals outstand over the existing research, Eigentrust [Kamvar et al., 2003] and P2Prep [Cornelli et al., 2002]. Eigentrust computes a single performance score for each peer, reflecting their past behavior in pairwise interactions. Although the protocol is distributed, it ultimately relies on a fixed set of trusted nodes at which it roots the computation of trust. On the other hand, P2Prep is designed for completely decentralized system. From the point of view of dissemination, and with the aim to make FIM systems decentralized, the strategy followed by P2Prep is directly applicable.

Based on this, we start the description of the IdMRep reputation protocol, which builds on the special features of FIM systems and introduces the desired P2P behavior. To apply the P2Prep dissemination, the FIM network can be modeled as a decentralized network based on the knowledge in the DTLs. We now detail the network model, as well as the protocol messages.

FIM Network Model

In order to achieve our goal to shift to a Peer-to-Peer behavior when establishing new federation relationships, entities should have a distributed way to find reputation information. For this purpose the data contained in the entities' DTLs can be used to define a new trust logic overlay.

The trust overlay is built in the following way: if a participating entity has a DTL entry for a specific entity in the FIM network, then there is a directed edge from the former to the latter. And this network model is what we call **"unstructured P2P based on DTL"** (see Figure 6.5). Since information will be only exchanged among trusted parties,

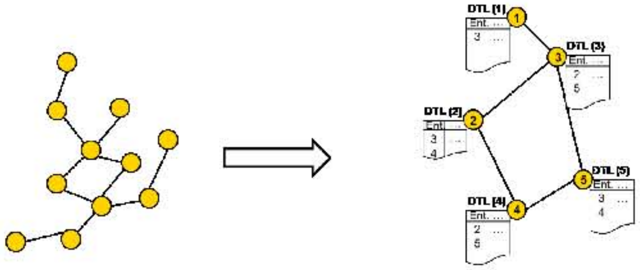


Figure 6.5: Unestructured P2P based on DTL

only trusted entries are used to build the logic overlay. On bootstrapping, DTLs are initialized based on the pre-configured trust relationships and agreements existing in each entity's TAL or BAL, and so the overlay is created based on these data. Then, the overlay will dynamically change according to the current state of the trust relationships based on entities' behavior, and on the entities joining/leaving the system. This improves the current and closed CoT model.

Over this DTL-based unstructured P2P model, we can now apply a dissemination protocol to gather reputation data. For this purpose, we define two new roles for entities participating in the reputation protocol, namely: `ReputationRequester` and `ReputationResponder`. Any IdP/SP in the network adopts these roles for asking and communicating reputation data, respectively.

Dissemination

In order to gather reputation data about an unknown entity the dissemination approach, which we call **"Query Flooding based on DTL"**, consists of broadcasting messages through the trust overlay. The protocol messages involved in this dissemination approach, conceptually depicted in Figure 6.6, are:

- `ReputationRequest`. This message is used to ask for reputation, it contains the following fields:
  - `Message_ID`: Message identification number
  - `ReputationRequester_ID`: The entity (SP or IdP) asking for reputation data
  - `Subject_ID`: The subject of the reputation or *"reputee"*, i.e., the entity (SP or IdP) whose reputation score is being calculated

- `Time to Live (TTL)`: Number of times the `ReputationRequest`is to be forwarded through the FIM network

- `Context (Cx)`: Reputation is associated to a context. In FIM, since the same entity can implement roles of SPs and IdP, these are considered contexts. That is, an entity can ask for the reputation of another as an SP or as an IdP. Appart from this role related contexts, a time context is also considered to allow the entities to ask for reputation data collected since a particular moment in time.

- `ReputationResponse`. This message is used to convey reputation data in reply to a `ReputationRequest`, it contains the following fields:

  - `Message_ID`: Message identification number (the same as in the request)

  - `ReputationResponder_ID`: The entity sending the reputation data (i.e., the *"reputor"*)

  - `Timestamp`: Time when the reputation message was issued

  - `Reputation data`: Associated to the entity identified by `Subject_ID`, regarding role contexts specified in the `Cx` field, and since the initial time specified in Cx

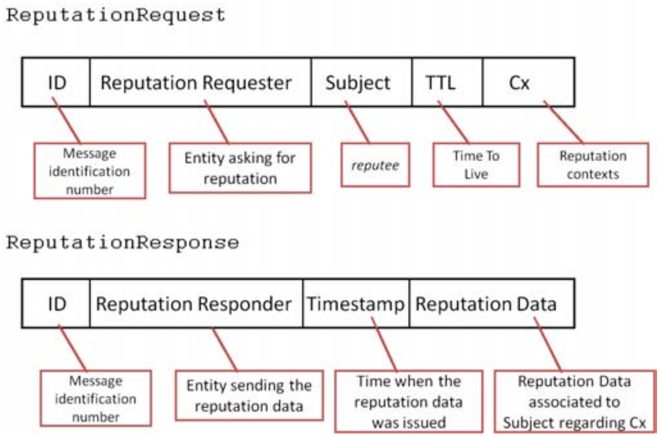

ReputationRequest
ID
Reputation Requester
Subject
TTL
Cx
Message identification number
Entity asking for reputation
reputee
Time To Live
Reputation contexts
ReputationResponse
ID
Reputation Responder
Timestamp
Reputation Data
Message identification number
Entity sending the reputation data
Time when the reputation data was issued
Reputation Data associated to Subject regarding Cx

Figure 6.6: IdMRep Protocol Messages

The envisioned dissemination approach in the IdMRep protocol is based on P2Prep, which relies in turn in the dissemination mechanisms in the P2P Gnutella network [Gnutella, 2003]. Accordingly, the above messages are used to gather reputation data by applying the following rules:

1. If an entity wants to gather reputation data about another one, then a `ReputationRequest` message is constructed and sent to the trusted entries in its DTL. A timer is set to wait for responses to this request

2. When receiving a message asking about the reputation of a specific subject (i.e., a `ReputationRequest` message), the entity must check its DTL to determine if there is an entry for the *reputee.* If an entry exists, then it must construct and send a response with the reputation data back to the requester (i.e., a `ReputationResponse` message)

3. An entity should forward incoming `ReputationRequest` messages to the trusted entities in its DTL, except to the one that delivered the incoming query and to the reputee (if it is in the DTL)

4. An entity receiving a message with the same `Message_ID`and reputee as one it has received before, must discard the message

5. After receiving all the `RequestResponse`messages (when the timer expires), the requesting entity must aggregate them to obtain a final global reputation value. The reputee is then added to the DTL with its associated computed reputation value.

`ReputationResponse` messages are sent directly to the requester instead of being routed backwards through the same way traversed by the `ReputationRequest`message. This is different as in P2Prep, which follows the Gnutella backwards routing. This choice of direct response means less overhead in terms of messages sent. Another difference with is in the primitives, which here are adapted to the specific semantics of FIM.

In Figures 6.7 and 6.8 we show a sample FIM network and the protocol sequence diagram for a particular transaction example in this network. More specifically, Service Provider SP1 wants to initiate a transaction with Identity Provider IdP2, which is unknown (no previous interaction). Thus, it executes the IdMRep protocol to obtain a T_R value and use it to compute a trustworthiness value for IdP2.

As far as the storage strategy is concerned, the DTL is used as the key element in order to maintain reputation related information.

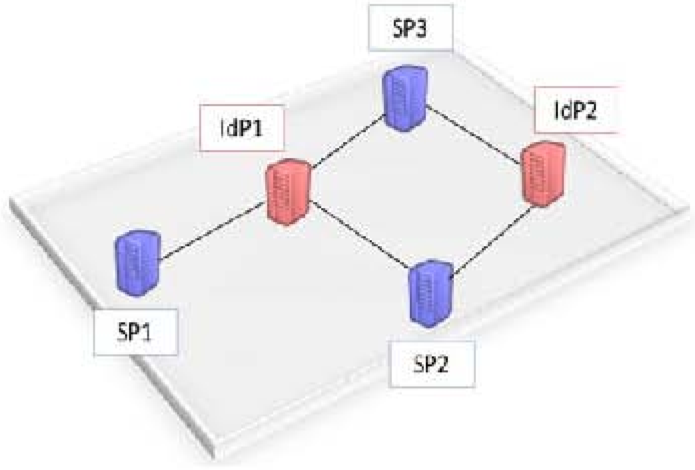
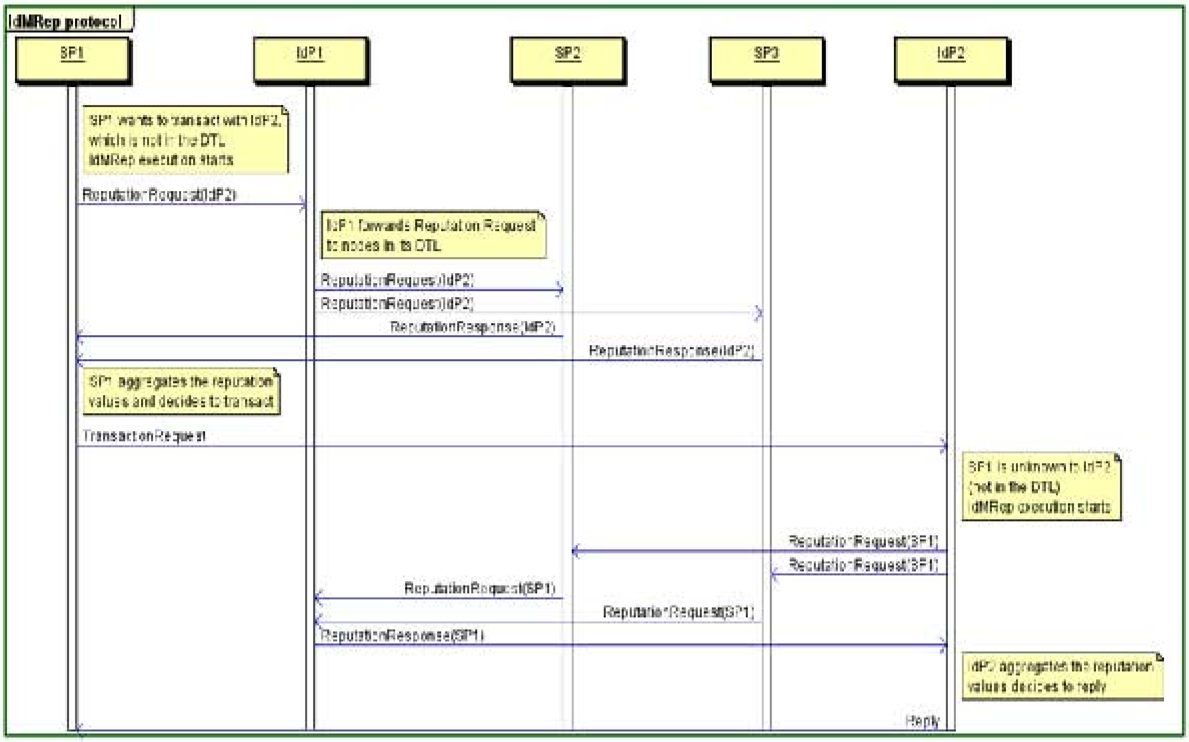


Figure 6.7: Sample FIM Network



SP1 wants to transact with IdP2, which is not in the DTL IdMRep execution starts

IdP1 forwards Reputation Request to nodes in its DTL

SP1 aggregates the reputation values and decides to transact

SP1 is unknown to IdP2 (not in the DTL) IdMRep execution starts

IdP2 aggregates the reputation values decides to reply

Figure 6.8: Sample IdMRep sequence diagram for the FIM network in 6.7

6.3.2. Further Details

Once the protocol model has been sketched, we provide an explanation of other complex aspects, namely the nature of the reputation data and its aggregation. Furthermore, issues such as security are discussed and some implementation guidelines are also given.

Reputation Scoring

Regarding the kind of information that is transmitted to convey reputation, there are different possibilities. Systems in the p2p reputation literature normally send its local trust value for the reputee computed with its own trust function based on interactions; or a single deterministic vote that express the opinion about the reputee. This vote can be binary, e.g., 0 if the repute is considered not reputable or 1 otherwise; or it can be a scaled integer, e.g., 1 to 10. But, this vote is also calculated based on the transaction history and applying the policies of the sender of the reputation.

These approaches are thus subjective in the sense that the reputation depends on the view of the reputation sender. In order to comply with our goal of making reputation more portable and less dependent on the reputor subjectivity, an approach that sends the raw history of transactions is more adequate. Thus entities can apply its own function over the transactional histories received in order to obtain a reputation value.

Any entity in the system must store transactional histories about the other entities with which they have interacted. The log system or repository stores for every transaction a tuple ⟨ `Time, EntityID, Role, Result` ⟩ , which contains the identifier of the entity with which the transaction was made, the role of that entity in the transaction, as well as de result (good or bad), all preceded by a time stamp. When a `ReputationRequest` is received, the fields in the `ReputationRequest` are used to filter the logs per entity and role, as well as to obtain the data from the origin time specified. Over these data, the number of bad and good transactions is calculated and packed into the `ReputationResponse`. Having defined how to convey reputation data and extract a reputation value from it, the next decision point in modeling the reputation system is how to aggregate the data from multiple sources. An easy and widely adopted approach is to use the arithmetic mean to obtain the average. But the choice will be application dependent.

According tho the above descriptions, the process to derive the final T_R value to be introduced in equation 6.2 is depicted in Figure 6.9. The first step consists of obtaining the transactional histories received from providers 1:N in the `ReputationResponse` messages. Then, step 2 involves the derivation of reputation values associated to each history (i.e., $T_R{}^1$:$T_R{}^N$). These values are obtained by means of the f_{evol} function used, which models the evolution based good and bad actions. Finally the third and last step consists of aggregating the T$_R{}^1$:$T_R{}^N$ values into a global T_R

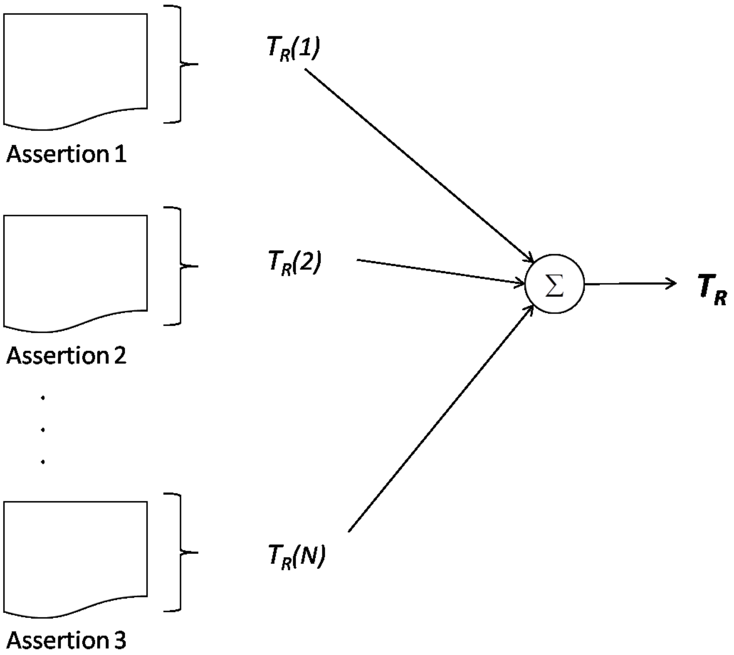


Figure 6.9: Derivation of T_R from *Reputation Assertions*

Integrating IdMRep in a SAML-based FIM network

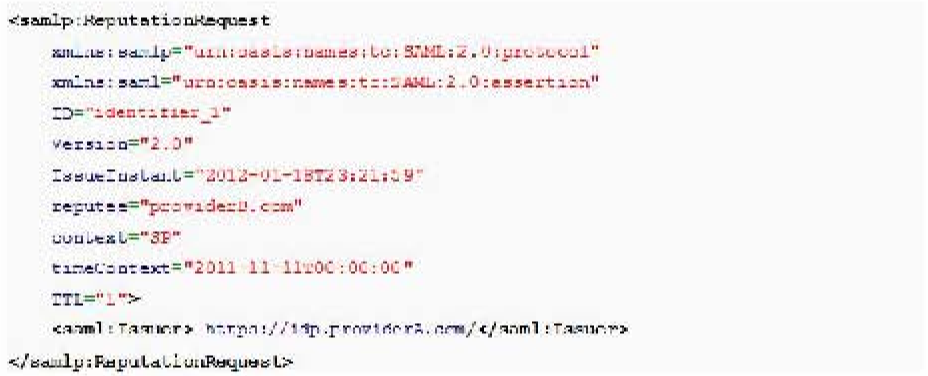
Normally, the primitives of a reputation protocol are implemented to be transported over the underlying protocols of the system where they are applied. For example, in reputation systems designed to improve file sharing in p2p networks, primitives are implemented over the p2p protocol used in that network. Similarly, since our reputation protocol is to be deployed in FIM networks, it is logical to implement it over FIM protocols (e.g., SAML, openID, etc.).

Since SAML is the best known standard FIM protocol, we detail how to implement the IdMRep protocol over it. SAML provides extension mechanisms that can be used for this purpose. Adding reputation support to SAML implies modifications to both assertions and protocols. The Reputation data in the IdMRep `ReputationResponse` primitive is thus expressed in the form of a SAML assertion. Regarding how the assertion is exchanged, it can be done as for a normal Assertion over the standard SAML protocols. More specifically, by using the SAML *Assertion Query and Request Protocol*, which is the basis to request or query for an Assertion. To achieve this, we define a new kind of assertion: the *Reputation Assertion*. Such Assertion contains a custom statement type, called `<ReputationStatement>` .

The structure of the *Reputation Assertion* has an initial part or `header`, whose content is the same that is defined in the standard assertions. This common section includes the assertion identifier, the names of the issuer and the subject, and information about the instant in which the assertion was issued. The XML tags are `<Assertion ID>`, `<Issuer>` , `<Subject>` and `<IssueInstant>`, respectively. And the content for this tags will be the value of `Message_ID`, `TimeStamp`, `ReputationResponder_ID`, and `Subject_ID`, defined in the IdMRep `ReputationResponse` primitive.

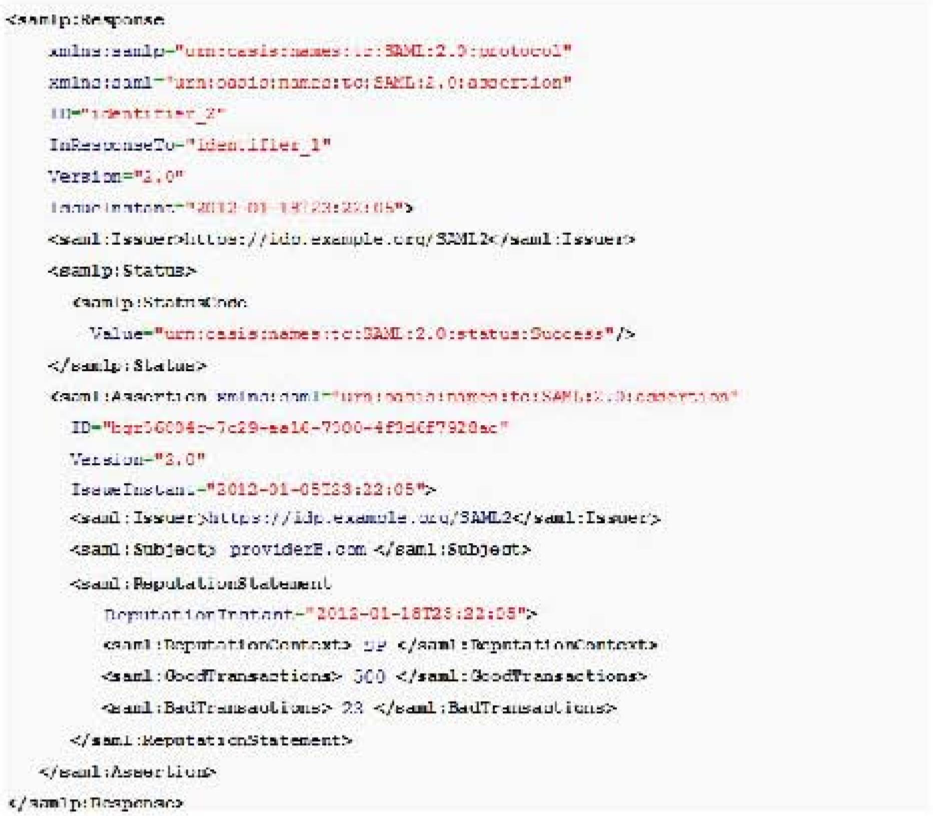
Apart from this information, the statement has a `body` section, which contains all the data related to the reputation metric. The tag `<ReputationStatement>` has been defined for this purpose. This tag includes the attribute `ReputationInstant`, to indicate the instant in time that is the origin of the history of transactions. Inside the `<ReputationStatement>` there are three more elements: `<ReputationContext>`, `<GoodTransactions>` and `<BadTransactions>`. The first element is used to indicate the role (SP, IdP) for what the reputation is expressed regarding the subject. The other two elements indicate the number of good and bad transactions that the subject performed with the issuer of the assertion. This assertion is conveyed as a SAML protocol `<Response>` .

On the other hand, in order to ask for a *Reputation Assertion*, we define a new element `<ReputationQuery>` used to make the query "*What assertions containing reputation statements are available for this subject?*" This element contains a `context` attribute that contains a string indicating the role (SP, IdP) for which the reputation about the subject is requested. An aditional `timeContext` attribute is also defined to express the initial time since the issuer wants reputation data about the subject. In order to express the subject, a `reputee` attribute is used. Furthermore, a `TTL` attribute is defined to indicate the horizon for request forwarding. Accordingly, Figures 6.10 and 6.11 show a sample SAML request message to ask for a *Reputation Assertion*, and a sample associated response message, respectively.



```
<samlp:ReputationRequest
    xmlns:samlp="urn:oasis:names:tc:SAML:2.0:protocol"
    xmlns:saml="urn:oasis:names:tc:SAML:2.0:assertion"
    ID="identifier_1"
    Version="2.0"
    IssueInstant="2012-01-18T23:21:59"
    reputee="providerB.com"
    context="SP"
    timeContext="2011-11-11T00:00:00"
    TTL="1">
    <saml:Issuer> https://idp.providerA.com/</saml:Issuer>
</samlp:ReputationRequest>
```

Figure 6.10: `ReputationRequest` over SAML



```
<samlp:Response
    xmlns:samlp="urn:oasis:names:tc:SAML:2.0:protocol"
    xmlns:saml="urn:oasis:names:tc:SAML:2.0:assertion"
    ID="identifier_2"
    InResponseTo="identifier_1"
    Version="2.0"
    IssueInstant="2012-01-18T23:22:05">
    <saml:Issuer>https://idp.example.org/SAML2</saml:Issuer>
    <samlp:Status>
      <samlp:StatusCode
        Value="urn:oasis:names:tc:SAML:2.0:status:Success"/>
    </samlp:Status>
    <saml:Assertion xmlns:saml="urn:oasis:names:tc:SAML:2.0:assertion"
      ID="bgr56004c-7c29-ea16-7300-4f3d6f7920ac"
      Version="2.0"
      IssueInstant="2012-01-05T23:22:05">
      <saml:Issuer>https://idp.example.org/SAML2</saml:Issuer>
      <saml:Subject> providerB.com </saml:Subject>
      <saml:ReputationStatement
          ReputationInstant="2012-01-18T23:22:05">
          <saml:ReputationContext> SP </saml:ReputationContext>
          <saml:GoodTransactions> 500 </saml:GoodTransactions>
          <saml:BadTransactions> 23 </saml:BadTransactions>
      </saml:ReputationStatement>
    </saml:Assertion>
</samlp:Response>
```

Figure 6.11: `ReputationResponse` over SAML

In the example of Figures 6.10 and 6.11, the requesting entity *ProviderA*, asks for the reputation of *ProviderB* acting as an *SP* since the *11st of November of 2011*. The response mesage contains the requested data inside a *Reputation Assertion*, indicating that the number of good actions since the requested date is 500, and the number of bad actions is

23.

For conveying the assertion, any binding defined in the SAML standard can be used as underlying transport mechanism, without requiring any further modification.

We have implemented a *proof-of-concept* prototype, which is explained in the validation chapter (Chapter 8) and provides further technical details on the implementation issues.

Furthermore, it is worth noting that the choice of implementing the IdMRep messages as SAML assertions is an advantage in the sense that most of the FIM protocols are able to convey SAML tokens. So messages can be used in other applications. When no SAML bearing mechanism is available, a translation service, to extract the assertion contents and translate into another token format can be used.

Security Issues

Regarding security, we rely on the SAML security mechanisms for the exchange of Assertions. These mechanisms are documented in the *"Security and Privacy considerations for the OASIS Security Assertion markup Language (SAML) v2.0"* [Maler et al., 2005]. Basically, by using SSL/TLS as transit protection protocol, confidentiality authentication and integrity in the communications between every pair of nodes is assured. It provides protection against eavesdropping attacks, message modification insertion or deletion and man-in-the-middle attacks. The selected cipher suite and the combination of transport layer security with message layer protection will lead to different security assurance levels, as it will be discussed in the risk chapter.

But these considerations are applicable only to the exchange of messages between two federated providers. However, the dissemination approach for querying and obtaining reputation data poses new security challenges. Similarly, the introduction of reputation makes the system vulnerable to attacks or flaws that are specific to reputation systems. We briefly discuss both kinds of security challenges:

- **Dissemination related security issues.**

  With the dissemination strategy described here, reputation data may be obtained from different sources. These sources may be direct trusted entities or totally unknown entities. For the case were the information comes from unknown providers, the credibility of the source may be questionable. Theoretically, the responding entity has been reached through a chain of trusted providers, but the requester does not know all the intermediate nodes (and their trustworthiness). Furthermore, any node in the forwarding chain knows about the poll and can use the data on the assertion with malicious purposes, e.g., submitting a bad fake reputation.

There are several ways to address these issues. One possibility is to adjust the TTL value to 1. If the horizon is the set of direct trusted neighbors (i.e., trusted entities in the DTL), the reputation data will come always from trusted sources. The inconvenient is that the network knowledge is limited to those peers located at a 1-hop distance.

Another mechanism is the modification of the protocol, so the `ReputationResponse` messages are routed back to the requester following the path of the `ReputationRequest` in such a way that every node in the chain adds information about his neighbor. In the end, the requesting node will have the reputation data of the subject for which the query was made, but also the length of the path through him, all the nodes traversed and the reputation of each of these nodes as seen by its direct neighbor in the chain. This approach has the benefit that nodes that route back the `ReputationResponse`, can store the reputation assertion in their DTLs and use it in future transactions. On the other hand, this type of routing increases the number of IdMRep messages required.

Another possibility is to incorporate a credibility measure that is assigned to unknown providers that send `ReputationResponse` messages and that allows to filter incorrect rates. For example, if the reputation data sent by an unknown provider is very different from that obtained from direct trusted providers, it can be considered as false and be discarded. Different credibility mechanisms are proposed in the literature that may be applicable.

- **Reputation related security issues.**

There are threats that particularly affect reputation systems. The authors in [Mármol and Pérez, 2009] do a good work in summarizing some of the most important and critical security threats that could be applied in reputation schemes designed for distributed systems.

The simplest threat is the existence of individual malicious peers. And the way of preventing such a misbehavior is by decreasing the level of trust or reputation of those participants who always provide bad services, categorizing them, therefore, as malicious peers. This threat is correctly addressed in our model, which is able to identify bad entities and isolate them (see Chapter 8).

Another kind of threats are related with the fact that entities may provide incorrect feedback to raise or decrease others reputation. The way of solving these problems is by introducing a notion of credibility of the entities in order to evaluate its trustworthiness also as reputors. This is not explicitly defined in our model but can be done easily by considering unfair reputation as a bad action. This will decrease the trustworthiness of the reputor and finally will be discarded also as a Service or Identity provider.

Most of the threats of reputation p2p systems come from two important features of these networks: there is usually no underlying security, i.e., cryptographic channels that allow the information to be exchanged confidentially; and there are usually non permanent identifiers assigned to participants. Not providing in transit security makes more feasible the injection of false reputation messages or modification/deletion of the messages. The non-persistent identifiers issue allows peers to perform bad actions until they are discovered and then disappear and start again with a new clean history (whitewashing). Furthermore, if the creation of identities is not very costly, the system is also susceptible of sybill attacks [Douceur, 2002].

In our case however, the cost of creating identities is high, since setting up a provider with FIM support is not an easy task. What is more, in FIM networks, providers have a well known identifier or URL. Behaving bad would be very costly to their reputation, which is difficult to recover again. They may end being excluded of the network, with the consequent economic losses.

To conclude, the improvements provided by including reputation in the FIM trust model are an increased flexibility and the tackle of uncertainty. The information gathered helps in acquiring knowledge about new providers facilitating cooperation with the trusted ones and preventing the selection of malicious ones. The incentive for good behavior is the winning of reputation, that will help in being admitted as cooperator by other providers.

6.4. Conclusions

In this chapter we have defined a computational trust model that captures the features of current FIM systems and introduces new dimensions to add flexibility and richness. With this, we move from the binary-based decision model currently used into a model that allows granularity. Trust is computed as a continuous numeric value and continually adjusted taking into account the behavior of the entities. An important part of the model is the inclusion of reputation data, which is a new research line in FIM that opens the door for further investigation. Here we just outlined a very simple protocol model that, far from reinventing the wheel, is based in the existing reputation protocols designed for p2p systems. We have defined the main features for the protocol to be applicable in FIM, the data that should be included and also how it can be implemented over a concrete specification (SAML). But further work can be done in determining the best possible protocol: studying the forwarding mechanisms that are more suitable, analyzing the threats and attacks, etc. All these potential studies are suggested as future research line in Chapter 9.

Furthermore, our proposal aims to combine trust with risk assessment in order to get a better informed decision-making procedure. Next chapter (Chapter 7) focuses on describing this procedure.

Chapter 7

Decision making scheme: To federate or not to federate

Although the future is uncertain, decisions have to be made, often in the light of incomplete and possibly incorrect information

Stephen Marsh

Contents

7.1. Chapter Overview

Not knowing an entity beforehand should not be a handicap to establish a federation. It is simply required that knowledge is gathered and, afterwards, make the decision based on this gathered knowledge in a dynamic fashion.

In the traditional FIM model entities were added and removed as entries in trust lists only as a result of explicit administrative action, reflecting changes to agreements with direct partners. The aim here is that these operations are performed automatically based on the *Decision Manager* output. This architectural component, outlined in Chapter 4, combines the trust and risk values associated to a transaction, and generates a final value that represents the decision. Thus, this Chapter concentrates on defining the operation of the *Decision Manager*, detailing the aggregation model followed.

7.2. Decision Manager: Rationale Design

7.2.1. Design Principles

In the proposed architecture for dynamic federation establishment and management, the *Decision Manager* is a key component. As introduced in Chapter 4, the function of this module is deciding whether to initiate or not a transaction with another entity, in case of being a requestor; or whether to respond or not to a transaction request. Thus, the goal in the Pre-Federation phase is to make a decision whether to federate or not with the other entity and to which extent; whereas in the Post-Federation phase, decisions are made about cooperating in particular transactions.

The inputs for this module (in terms of assurance), are the trustworthiness value of the other entity, the risk associated to the transaction, and the internal policies that will be used to govern the decisions. Then, the decision procedure consists of aggregating two metrics in order to obtain a final meaningful figure that can be compared to a decision threshold. We denote this final aggregated metric as **decision trust**, and its components are depicted in the image of Figure 7.1.

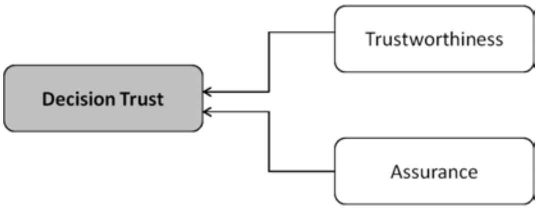
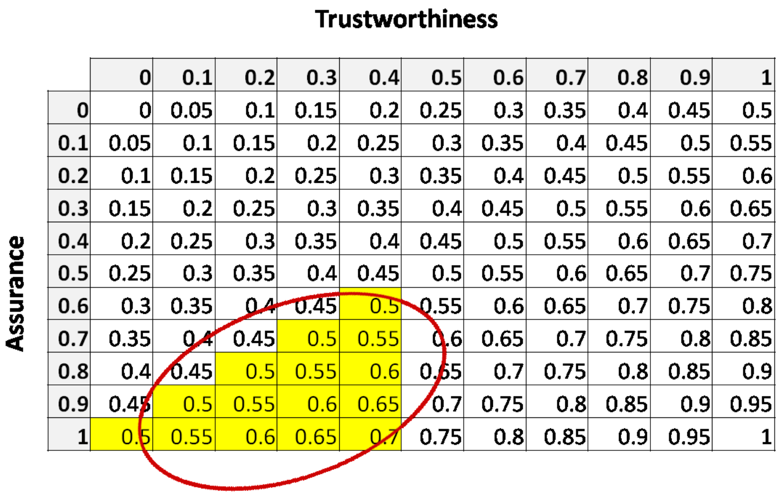


Figure 7.1: Dimensions of the decision trust metric

The process for obtaining the trustworthiness metric was detailed in Chapter 6. Similarly, the process for obtaining the assurance metric, which is directly proportional to the existing risk, was developed in Chapter 5. In this latter case, the MAUT theory was used to aggregate the several dimensions of risk, because this is a useful technique to combine criteria of different nature.

Since the dimensions that have to be aggregated for obtaining the decision trust value are also criteria of different nature, the application of MAUT for the aggregation of trustworthiness and assurance is reasonable. MAUT models permit to allocate different kind of aggregation functions, being additive linear functions the simplest ones and the most usually adopted. Yet, simple models such as weighted sums are not always sufficient. One assumption that must be held to apply linear aggregation in MAUT is that criteria are independent. Classically, if *mutual preferential independence* can further be assumed, the weighted arithmetic mean is used. That was the case in the aggregation of risk dimensions in Chapter 5. But this independence property is not present in the risk-trust relationship. For example, normally the higher the trust, the higher the risk we are willing to assume (see section 3.4.2).

To demonstrate the unsuitability of the basic MAUT aggregation, Figure 7.2 shows the matrix of results after aggregating trustworthiness and assurance by applying the most common function: the arithmetic mean. The surface for the decision trust applying this aggregator is depicted in Figure 7.3.



Trustworthiness

Assurance	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
0	0	0.05	0.1	0.15	0.2	0.25	0.3	0.35	0.4	0.45	0.5
0.1	0.05	0.1	0.15	0.2	0.25	0.3	0.35	0.4	0.45	0.5	0.55
0.2	0.1	0.15	0.2	0.25	0.3	0.35	0.4	0.45	0.5	0.55	0.6
0.3	0.15	0.2	0.25	0.3	0.35	0.4	0.45	0.5	0.55	0.6	0.65
0.4	0.2	0.25	0.3	0.35	0.4	0.45	0.5	0.55	0.6	0.65	0.7
0.5	0.25	0.3	0.35	0.4	0.45	0.5	0.55	0.6	0.65	0.7	0.75
0.6	0.3	0.35	0.4	0.45	0.5	0.55	0.6	0.65	0.7	0.75	0.8
0.7	0.35	0.4	0.45	0.5	0.55	0.6	0.65	0.7	0.75	0.8	0.85
0.8	0.4	0.45	0.5	0.55	0.6	0.65	0.7	0.75	0.8	0.85	0.9
0.9	0.45	0.5	0.55	0.6	0.65	0.7	0.75	0.8	0.85	0.9	0.95
1	0.5	0.55	0.6	0.65	0.7	0.75	0.8	0.85	0.9	0.95	1

Figure 7.2: Aggregation of trustworthiness and assurance by applying the arithmetic mean

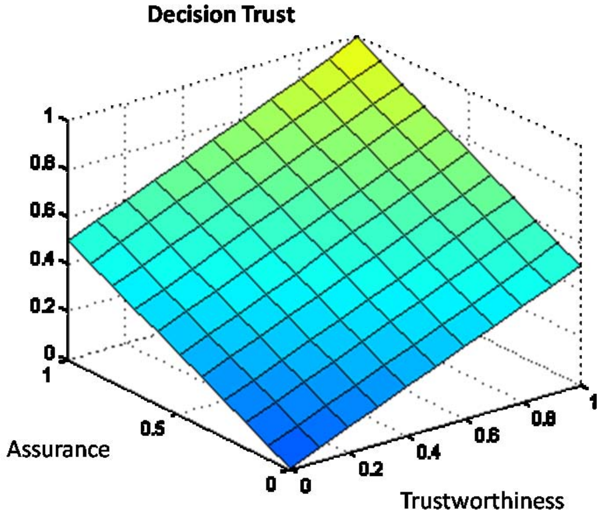


Figure 7.3: Decision Trust surface for aggregation using the arithmetic mean

As we can observe in the decision trust values represented in Figure 7.2, there are cases in which very low values of trustworthiness combined with high values of assurance lead to final acceptable values, because of the compensation effect of the aggregator. A 0.5 value can be e.g., obtained by the aggregation of maximum assurance and minimum trustworthiness, as well as for minimum assurance and maximum trustworthiness, and also for intermediate values. The final result does not provide these semantics and the difference is important. If combinations of extreme values were equally preferred, then the property of *additive independence* would hold, and the arithmetic mean aggregator would fit well.

Thus, we can see that an operator such as the weighed mean cannot express any interaction between criteria. Important behavioral features that should be fulfilled by the decisor are not captured. More specifically, to contextualize the problem, the features that the *Decision Manager* must fulfill are detailed below:

- Combine a trustworthiness value in the continuous range [0,1] with an assurance value in the continuous range [A_{min}, 1]. The 0 in the trustworthiness scale means distrust, 0.5 is ignorance and 1 is full trust. In the case of the assurance metric, the scale goes from a minimum assurance value A_{min} until 1, which means full assurance. It should be noted that, following the methodology in Chapter 5, the assurance assigned to transactions which do not fulfill the minimum requirements is equal to 0. Thus, all the values below this A_{min} threshold are assimilated to 0 and can be filtered before performing the final aggregation. That is, transactions with assurance under the A_{min} level are directly rejected. For the cases above A_{min}, assurance is combined with trustworthiness to make the decision. Assuming then that the assurance satisfies the minimum, the combination of this dimension with the trustworthiness must be performed in such a way that:

- If the level of trustworthiness is low, i.e., the entity is considered untrustworthy, then the decision must be to not cooperate.
- If the level of trustworthiness is high, i.e., the entity is considered trustworthy, then the decision must be to cooperate. However, depending on the assurance, value the extent of cooperation may be different.
- If the level of trustworthiness is dubious (e.g., close to 0.5), there is more uncertainty on whether the entity will behave good or bad. In these case, cooperation should be only permitted if the assurance level is high (i.e., low risk). If the assurance is low, even if acceptable, transactions should be avoided.

- Provide an output value, the trust decision metric, in the range [0,1]. A cooperation threshold (c_{th}) must be defined in the policies, e.g., typically 0.5.

These premises are the design principles for our *Decision Manager*. The decision model has indeed to be rich enough to model the decisional behaviors explained above. Since linear aggregation functions are not satisfactory, we need to define a different aggregation model. The desired decision surface, expressed in a matrix form, is shown in Figure 7.4.

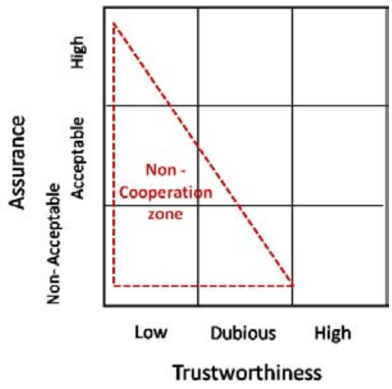


Figure 7.4: Matrix for the decision surface

7.2.2. First Approach: Hand-crafted Aggregation

Thinking on the above premises, we can design a hand-crafted aggregation function that meets them. For example, assuming that:

a) trustworthiness is considered low under 0.4, dubious between 0.4 and 0.6, and high over 0.6;

b) minimum assurance is 0.2, acceptable assurance is from 0.2 to 0.8, and high over 0.8;

c) the c_{th} is 0.5, i.e., a decision trust value equal or over 0.5 means a positive decision,

we have the cooperation map depicted in Figure 7.5.

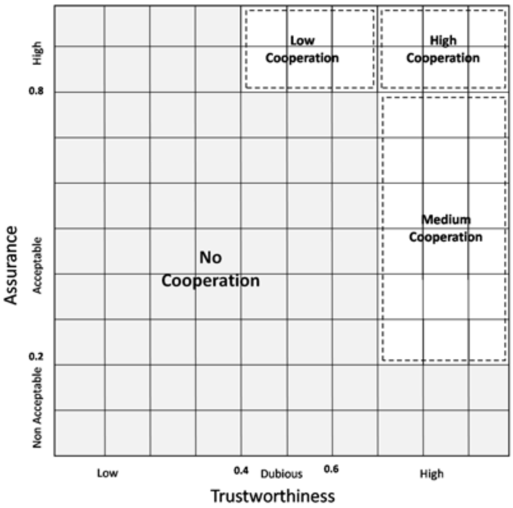


Figure 7.5: Cooperation map for *Decision Manager* with fixed thresholds

In order to obtain coherent output values, a piecewise function can be applied that guarantees values lower than 0.5 in the “no cooperation zone”, and values equal or over this threshold in the “cooperation zone”. Formulas 7.1 and 7.2 fulfill the desired behavior. Its associated surfaces are depicted in Figure 7.6 for equation 7.1, and Figure 7.7 for equation 7.2.

$$DT = \begin{cases} T \cdot A \;\; if \;\; T < 0.4 \;\; or \;\; if \;\; 0.4 \geq T \leq 0.6, A \leq 0.8 \\ 0.5 \cdot T + 0.5 \cdot A \;\; otherwise \end{cases} \tag{7.1}$$

$$DT = \begin{cases} 0 \;\; if \;\; T < 0.4 \;\; or \;\; if \;\; 0.4 \geq T \leq 0.6, A \leq 0.8 \\ 0.5 \cdot T + 0.5 \cdot A \;\; otherwise \end{cases} \tag{7.2}$$

where DT is decision trust, T is trustworthiness and A is Assurance.

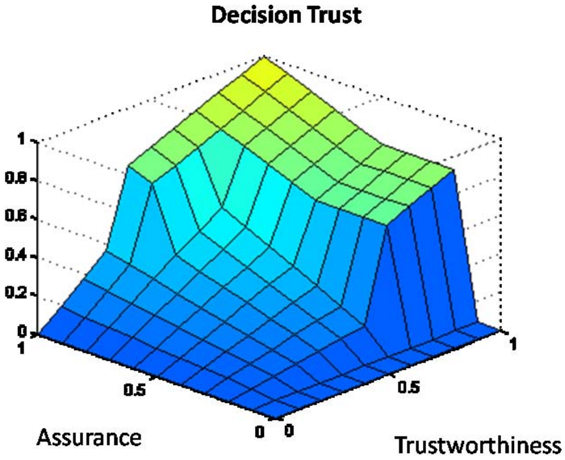


Figure 7.6: Decision trust surface for aggregation with equation 7.1

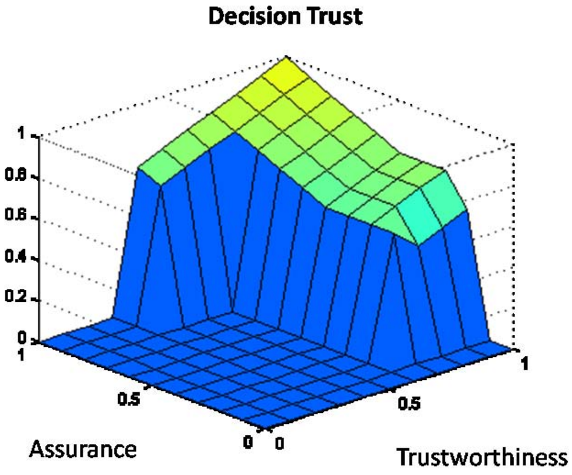


Figure 7.7: Decision trust surface for aggregation with equation 7.2

However, despite the above hand-crafted functions are able to model the desired features of the decisor, some limitations exist. According to the propositions stated above for the operation of the decision making module, the decision trust varies according to the values of assurance and trustworthiness, which are described in a qualitative fashion: low, acceptable, etc. But the numeric ranges for these categories will depend on the view of the evaluating entity, and so are subjective. For example, a 0.4 value for trustworthiness may be dubious from the point of view of a particular entity, or may be considered low from the point of view of another entity. Furthermore, when using a hand-crafted function a change on the selected thresholds may lead to incoherent results and so the aggregation function needs to be redefined.

7.2.3. Second Approach: Fuzzy Aggregation

To capture the vagueness and imprecision of the operation propositions, as well as the existent subjectivity, a richer model needs to be defined. Furthermore, the model has to be parametrizable, so the aggregation function does not require changes when the ranges of the categories vary.

In this sense, fuzzy aggregation [Beliakov and Warren, 2001] techniques are appropriate to be applied. Fuzzy logic provides a mathematical formalism for a unified treatment of vagueness and imprecision that are ever present in decision support and expert systems in many areas. In fact, in the specific context of risk-trust relationship, Manchala [Manchala, 2000] proposes to express the interaction of both concepts by using a fuzzy logic trust matrix (similar to the representations in Figures 7.4 and 7.5). This work points out that the fuzzy trust matrix could be replaced by a set of fuzzy membership functions that could be useful in reasoning. Thus, we will base on this theory to build the model for our *Decision Manager* system.

In fuzzy set theory (FST), membership functions of fuzzy sets play the role similar to the utility functions in MAUT (the role of degrees of preference). Consequently, we have just to design the system including the membership/utility functions and perform aggregation.

More specifically, Mandani fuzzy inference systems (FIS) are widely used in particular for decision support applications since they are intuitive and easy to interpret. In order to construct a Mandani FIS, the following steps are performed: (1) Fuzzification, which comprises definition of the input and output parameters and its associated levels or linguistic labels; (2) Definition of *If-Then Rules*, which will be used for reasoning about input values and obtain the output; and (3) Deffuzification, which refers to the calculation of a single output number after the rules are applied.

The first step allows us to obtain the MAUT utility functions to be aggregated. In our case, the two **input variables** for the FIS are trustworthiness and assurance. For the input variable trustworthiness, we define three linguistic labels, namely: *untrusted*, *dubious*, and *trusted.* The quantitative values assigned to each label depend on two thresholds defined in the decision making policies, namely:

- Dubious distrust threshold (d_{th}). For values of trustworthiness below this threshold an entity is considered untrustworthy.
- Dubious trust threshold (t_{th}). For values of trustworthiness over this threshold an entity is considered trustworthy

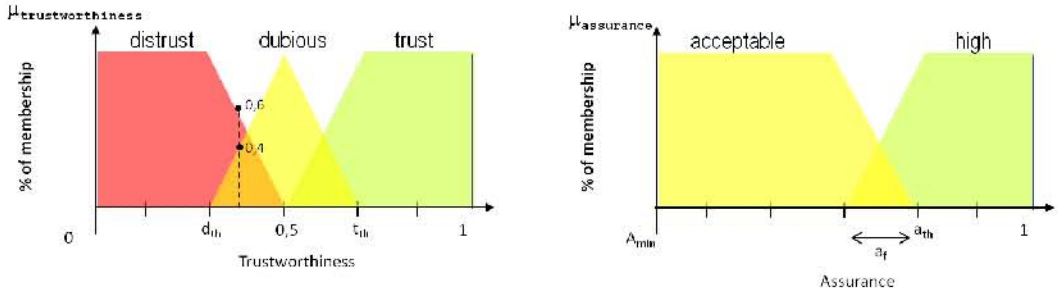
For the input variable assurance, we define two labels: *acceptable* and *high.* The threshold that separates both categories is called high assurance threshold (a_{th}), and must be also defined in the policies. For values over this threshold, assurance is considered high.

These features of the input variables are depicted in Figure 7.8. In the trustworthiness graph, it can be observed that thresholds d_{th} and t_{th} determine the range of the linguistic categories. Furthermore, there is a zone where values overlap. For example, a particular value between d_{th} and 0.5 can be interpreted as distrust or as dubious. The membership function (μ) indicates how much the value belongs to each category, e.g., the trustworthiness value associated to an entity can be a 60% distrust and a 40% dubious. That is, it belongs to both categories with different degrees of membership. This feature of fuzzy logic allows to capture subjectivity.

In the assurance graph, the a_{th} threshold determines from which point the assurance can be considered 100% high. In order to introduce a degree of fuzziness, such as the dubious range in trustworthiness, we define an assurance flexibility index a_f as the range in which the assurance could be considered both high and acceptable.

For the **output variable** decision trust, we define four labels: *non-cooperation*, *low*, *medium* and *high.* The quantitative values assigned to each label depend on a set of thresholds defined in the decision making policies. These thresholds are:

- Cooperation threshold (c_{th}). For values of decision trust below this threshold the decision to cooperate will be negative.
- Low Cooperation threshold (lc_{th}). For values of trustworthiness over this threshold the decision to cooperate will be positive, and the extent of the cooperation low.



(a) Fuzzy trustworthiness

(b) Fuzzy Assurance

Figure 7.8: Trustworthiness (a), and assurance(b)

- Medium Cooperation threshold (mc_{th}). For values of assurance over this threshold the decision to cooperate will be positive and the extent of the cooperation medium.
- High Cooperation threshold (hc_{th}). For values of assurance over this threshold the decision to cooperate will be positive and the extent of the cooperation high.

Apart from these thresholds, we also define the flexibility indexes lc_f, mc_f, and hc_f, to define fuzzy ranges in which categories overlap. The decision trust output variable is depicted in Figure 7.9

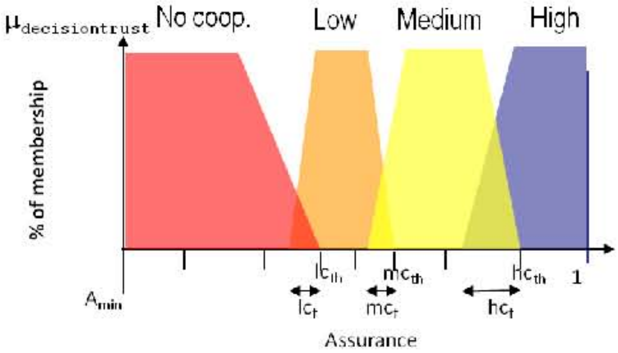


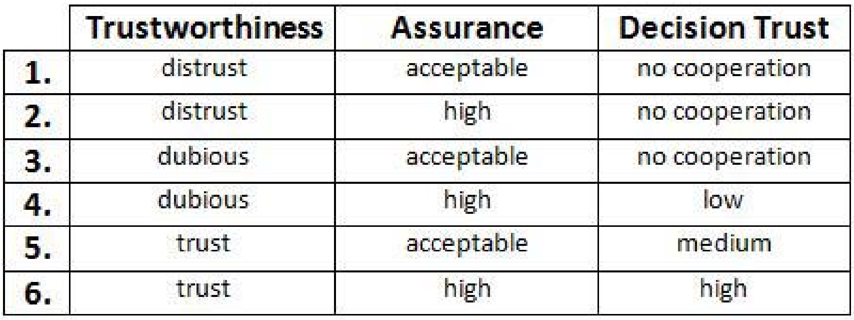
Figure 7.9: Fuzzy decision trust

Having defining the inputs and the output, the next step is the definition of the *If-Then Rules* for reasoning in order to obtain an aggregated output. A fuzzy rule is defined as a conditional statement in the form:

“**IF** x is A **AND** y is B, **THEN** z is C”

where x and y are the linguistic input variables; z is the linguistic output variable; and A, B and C are the linguistic values of the variables, respectively. Different logical operators, or combinations of them, can be used to formulate the rules (e.g., disjunction OR).

The complete set of rules we have defined to model the desired component for the *Decision Manager* are presented in Figure 7.10. All of them are formulated using the conjunction operator AND. To give an example, the logical formulation of the first rule would be: "**IF** *trustworthiness* is *distrust* **AND** *assurance* is *acceptable*, **THEN** *decision trust* is *no-cooperation*"



	Trustworthiness	Assurance	Decision Trust
1.	distrust	acceptable	no cooperation
2.	distrust	high	no cooperation
3.	dubious	acceptable	no cooperation
4.	dubious	high	low
5.	trust	acceptable	medium
6.	trust	high	high

Figure 7.10: Fuzzy inference rules (knowledge base for the *Decision Manager*)

By applying the rules in in Figure 7.10 and defuzzifying the output to get a quantitative value, the decision trust is obtained. The aggregation process is graphically shown in Figure 7.12. There are various mathematical operations underlying the whole process: maximum, minimum, products, t-norms, etc. More details can be found in [Fodor and Roubens, 1994].

Now that the whole aggregation model is defined, we show in Figure 7.11 the decision surface obtained when assigning the thresholds to the same values assumed in the example described in 7.2.2. That is t$_{th}$= 0.6 , d$_{th}$ = 0.4 for the trustworthiness; and a$_{th}$= 0.8 for the assurance. The assurance includes a flexibility index a$_{f}$=0.05.

As it can be noticed from the image in Figure 7.11, the decision surface is similar to that obtained when applying the hand-crafted functions in equations equation 7.1 and 7.2. However it has the advantage that is parametrizable, so thresholds can be changed without the need to redefining the system. This FIS-based system properly models the richness

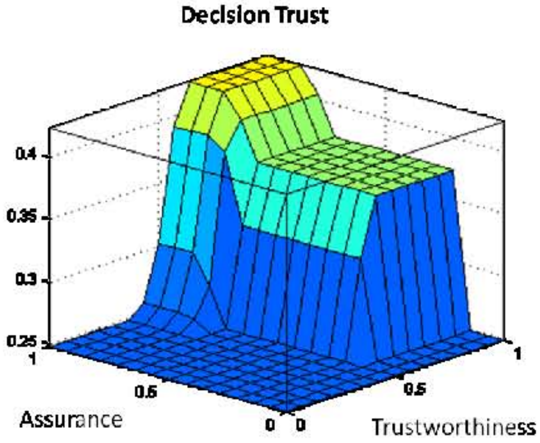


Figure 7.11: Decision trust surface with Mamdani FIS based aggregation

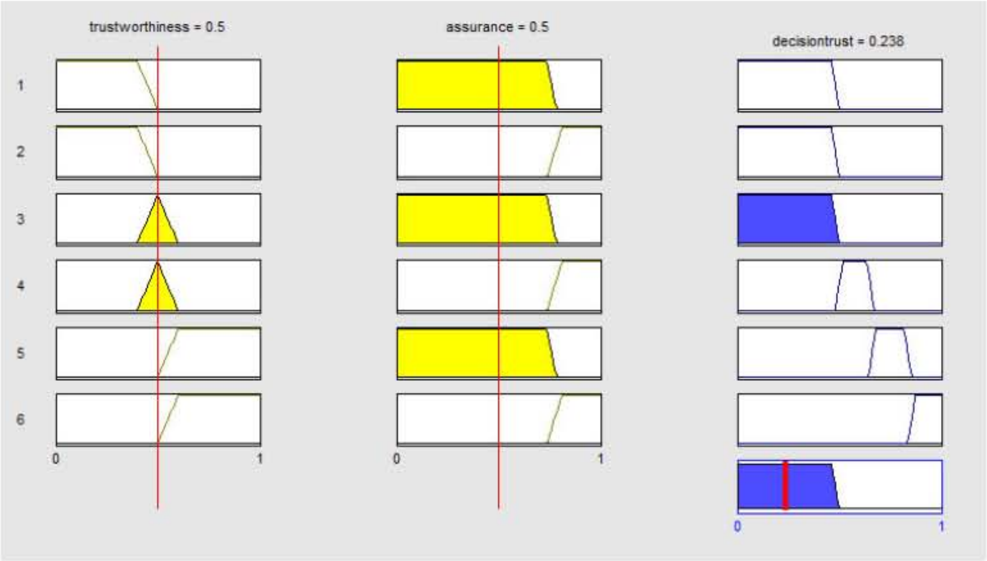


Figure 7.12: Mamdani FIS based aggregation

of the *Decision Manager* and allows to obtain an output mapped in multiple categories, which are useful to constrain the extent of the cooperation (e.g., using different SLAs for each category).

7.3. Conclusions

The acceptance of risk and the means, via trust, to cope with and assimilate it into decisions, enables humans to exist in the complex society which is around us [Luhmann, 1979]. The formalization of these social concepts into mathematical models make them usable in technically-mediated interactions, so they can be performed on the basis of the same concepts we ourselves rely on. Thus, with the introduction of these mechanisms in FIM, decisions can be made in an automated and more dynamic fashion.

Based on these premises, this chapter described how to combine trust and risk to output a decision to cooperate. We started the definition procedure by trying the simplest aggregation model, i.e., a lineal additive function. However it resulted to be limited due to the complex relationship between the trust and risk concepts. The next step consisted on the definition of an ad-hoc function, that leads to the desired results but is limited in regard with flexibility, i.e., if the entity policies change, then the function must be redefined. Finally, we evolved towards a fuzzy-based aggregation system, that is parametrizable, flexible, allows to model complexities in the trust-risk relationship and captures the subjectivity of entities.

Chapter 8

Validation

Contents

8.1. Chapter Overview

This chapter is dedicated to cover the validation of the ideas presented in this thesis. Its central part is organized in three main blocks that are related to each of the main contributions. Firstly, section 8.2 presents a formal analytical validation of the mathematical risk model and the adjustment of its parameters. Secondly, section 8.3 describes a simulation model used for the evaluation of the benefits of including reputation data to build trust in a FIM network. Thirdly, section 8.4 develops the validation of the proposed architecture through the implementation a proof-of-concept prototype.

Finally, section 8.5 concludes by remarking the main results derived from all the the validation tests performed, as well as identifying which aspects still need to be covered.

8.2. Risk Model Validation: Aggregation Tests, Examples and Discussion

This section is dedicated to validate the risk model and assessment procedure proposed in Chapter 5. Validation, based on NIST Special Publication 500-238 definitions [Wallace et al., 1996], is understood as the process of checking whether the proposed solution satisfies its expected requirements. That is, we conduct validation with the aim to demonstrate that the model works in conformance to the associated principle guidelines, that the output is correct. More specifically, we test how the model is capable of handling the set of quantitative metrics defined and use them as an input to generate the associated risk value. We show that the final risk value:

- is relative to the perception, assets and needs of the provider that is making the evaluation
- provides information on the assurance level coverage
- can be used in decision making. The final value allows to discard those entities that do not satisfy minimum requirements; and also to make a comparative ranking of entities when there are several options available

We go through the validation process starting with aggregation tests at the lowest level in the risk hierarchy (level L4) and at the immediate upper level (level L3), and then we provide a complete aggregation example involving the whole Pre-Federation risk branch. It is to mention that through the tests we refine the model and complete its definition. Thus, recommendations are given below on how to implement aggregation.

8.2.1. Test 1: Aggregation of L4 Metrics

Taking the Pre-Federation branch of the taxonomy for risk assessment metrics, we have the hierarchy in Figure 8.1.

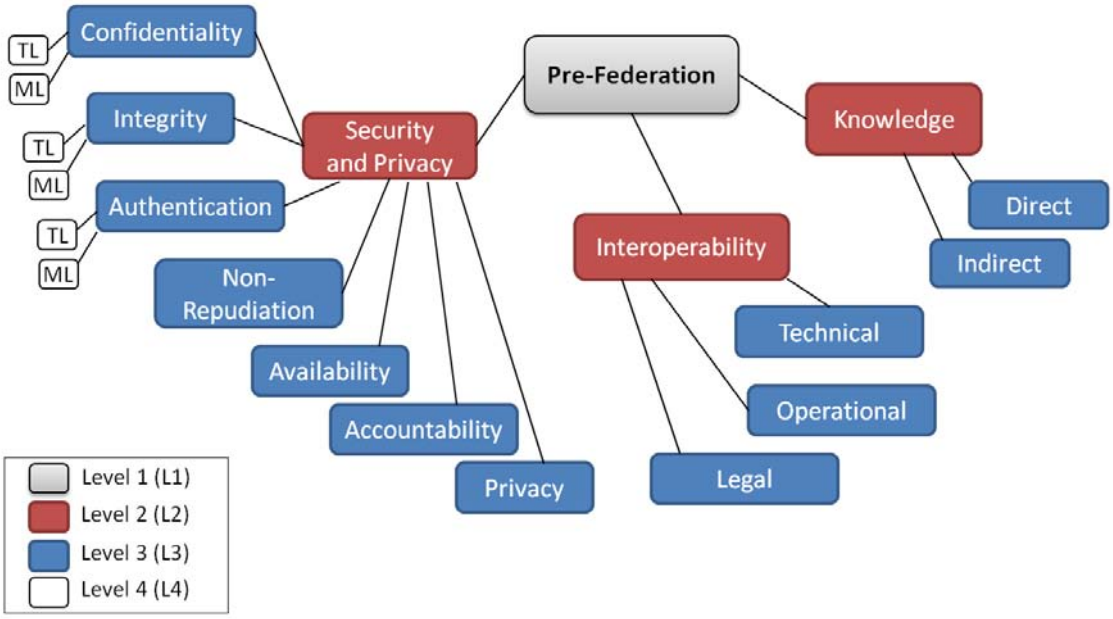


Figure 8.1: Pre-Federation hierarchy

The first step for obtaining the global Pre-Federation assurance of a provider A under evaluation is applying MAUT to aggregate the criteria located at level 4 and reduce the problem from a 4-level to a 3-level hierarchy. The proposed set of security related criteria includes three cases where aggregation must be performed: confidentiality, integrity and authentication. All of them are to be calculated by combining the assurance provided both at transport and at message layers. To analyze the aggregation at this level we take as example the confidentiality dimension, but the procedure would be the same for integrity and authentication. According to the hierarchical aggregation formula in 5.5, the global Confidentiality Assurance score is obtained as:

$$Agg^{3,1}(A) = \sum_{i}^{n} \omega_i^{4,1} \cdot g_i^{4,1}(A)^T = WV^{4,1} \cdot SV^{4,1^T}(A)$$

$$= WV^{4,1} \cdot [|CONF_{TL}(A)|, |CONF_{ML}(A)]^T$$

Where |CONF$_{TL}$(A)| and |CONF$_{ML}$(A)| represent the normalized values for the confidentiality metrics at transport and message layers respectively. Table 8.1 shows the aggregated values for all the possible sub-criteria combinations under the confidentiality dimension applying different weights.

CONF$_{TL}$	**$g_1{}^{4,1}$**	**CONF$_{ML}$**	**$g_2{}^{4,1}$**	**weights [0.5,0.5]**	**weights [1,0]**	**weights [0,1]**
0	0	0	0	0	0	0
1	1/3	0	0	1/6	1/3	0
2	2/3	0	0	1/3	2/3	0
3	1	0	0	1/2	1	0
0	0	1	1/3	1/6	0	1/3
1	1/3	1	1/3	1/3	1/3	1/3
2	2/3	1	1/3	1/2	2/3	1/3
3	1	1	1/3	2/3	1	1/3
0	0	2	2/3	1/3	0	2/3
1	1/3	2	2/3	1/2	1/3	2/3
2	2/3	2	2/3	2/3	2/3	2/3
3	1	2	2/3	5/6	1	2/3
0	0	3	1	1/2	0	1
1	1/3	3	1	2/3	1/3	1
2	2/3	3	1	5/6	2/3	1
3	1	3	1	1	1	1

Table 8.1: Confidentiality Assurance as a result of aggregating normalized CONF$_{\mathrm{TL}}$ and CONF$_{\mathrm{ML}}$ using different weights in the aggregation

As shown in Table 8.1, the aggregated results present important variations depending on the chosen weights. Applying extreme weights (0 or 1) leads to take into consideration only one of the criteria. On the other hand, for the case of using equal weights (equivalent to an arithmetic mean), the result is a balanced combination of the criteria. However, meaningful differences in the partial contributory factors may be hidden in the final value. For example, the final confidentiality assurance value using equal weights is the same for an entity that provides the maximum assurance on the transport layer and no assurance on the message layer, but also for an entity that provides no assurance at transport and maximum assurance at message. One could think that the higher the security assurance the better, but a higher assurance is not representative since it may contain low values in specific provisions that can be important for the party evaluating risks.

These facts point out the need of improving the aggregated metrics adapting the weights in order to convey more information about the partial contributory factors and allow better informed decisions. The weights must be then calculated according to the interests of the evaluating party. Our approach for their derivation is another contribution to the modeling phase and it is based on the minimum risk requirements of the evaluating party.

We are building a risk model but so far we have just talked about assurance. The assurance values show only partial risk information. For example, confidentiality risk at transport layer is inversely proportional to the confidentiality assurance provided at this layer ($\mathrm{CONF_{TL}}$). The more the confidentiality assurance, the less the probability ($\mathrm{P_{CONF_{TL}}}$) of incurring in risks related to confidentiality ($\mathrm{P_{CONF_{TL}}}\ \alpha\ 1/\mathrm{CONF_{TL}}$). But, on the other hand, every organization will have different sensitivities to different attacks depending on the value of the assets under its control and other contextual features. So risk is also proportional to impacts (I). In this case, it means that $\mathrm{RISK\text{-}CONF_{TL}} = \mathrm{P_{CONF_{TL}}}$ x $\mathrm{I_{CONF_{TL}}}$. Based on this, a qualitative scale for $\mathrm{P_{CONF_{TL}}}$ that reflects the inverse relation with the assurance is shown in Table 8.2.

CONF$_{TL}$	**P**$_{\mathbf{CONF_{TL}}}$
High (H)	Low
Medium (M)	Medium
Low (L)	High
None (N)	Very High (VH)

Table 8.2: Inverse relationship between probability and assurance

Then, having the $\mathrm{CONF_{TL}}$ value, the confidentiality risk can be obtained by means of a risk matrix including both impact and probability dimensions, as shown in the example of Table 8.3. This kind of matrices is to be defined in the local policies of entities according to their risk criteria.

	I High (100)	**I Medium (50)**	**I Low (10)**
P Very High (1)	VH (100)	H (50)	L (10)
P High (0.9)	H (90)	M (45)	L (9)
P Medium (0.5)	H (50)	M (25)	L (5)
P Low (0.1)	L (10)	L (5)	L (1)

Table 8.3: Example of risk matrix

The numeric values for probabilities and impacts in the risk matrices are to be defined also in local policies. Here, for the sake of illustrating the methodology with a numeric example, we use values similar to the ones in the NIST risk assessment methodology [Stoneburner et al., 2002]. Normally, when evaluating an entity for cooperation it is interesting that it satisfies some minimum requirements in regard to different criteria. For example, in regard to confidentiality a local policy may say that, as a minimum requisite, low assurance must be provided at both transport and message layers. However, it is impossible to tell if minimum requirements are met from the global aggregated values presented in Table 8.3. With the idea to solve this problem we refine the aggregated metrics. For this purpose, the following notation is used:

- $RV^{k,j}$ is a *Reference Vector* such that $RV^{k,j}(i)$ contains the minimum required value for the *i-th* assurance metric in SV^{k,j}.
- $|\cup\ SV^{k,j}|$ the number of metrics in the score vector of a provider ($SV^{k,j}$) that are greater or equal than the minimum required value (i.e., the number of metrics that fulfill $SV^{k,j}(i) \geq RV^{k,j}(i)$)

Based on the above notation and inspired by the compliance metric proposed in [Luna et al., 2011], let us apply their concept to the FIM environment and define the *Assurance Compliance Index* (ACI) as in expression 8.1:

$$ACI^{k,j} = \begin{cases} 1 \ \ if \ \ SV^{k,j}(i) \geq RV^{k,j}(i) \ \ \forall \ \ i \\ \dfrac{|\cup SV^{k,j}|}{n} \ \ otherwise \end{cases} \tag{8.1}$$

Thus, an ACI equal to 1 means that the minimum requirements are satisfied; otherwise it gives an idea of the requirements coverage. Based on this compliance index, we define the *Constrained Aggregated Assurance* in expression 8.2, which improves the original aggregation in 5.5, making it more meaningful:

$$CAgg^{k,j} = \begin{cases} Agg^{k,j} \ \ if \ \ ACI^{k,j} = 1 \\ 0 \ \ if \ \ ACI^{k,j} \neq 1 \end{cases} \tag{8.2}$$

Based on this, and assuming that entities have risk matrices for every criterion, the following steps must be followed to derive the minimum assurance requirements that fulfill risk needs and construct the reference vector $RV^{(k,j)}$:

1. consult the local policies to determine the impacts and maximum assumable risk for each criterion;

2. use the risk matrices to identify the corresponding desired probabilities of incurring in risk using the impact-maximum risk pairs; and

3. derive the required minimum assurance level for each criterion based on the desired probabilities

The reference vector, apart from showing the minimum requirements, it also gives a notion of which criteria are more important. It is rational to assume that if the requirements on a particular dimension are higher with respect to the other dimensions, then the relative importance between them is high. Based on this, and following MAUT recommendations for weighing, we rate each criterion with the minimum required assurance and divide by the summation of all the elements in the reference vector to get the corresponding weight. Mathematically, expression 8.3 shows how to obtain the weights:

$$WV^{k,j} = \frac{RV^{k,j}}{\sum_i^n RV^{k,j}(i)} \tag{8.3}$$

Thus, by deriving the weights as in 8.3 and applying formula 8.2, the metrics at level 4 are aggregated and the hierarchy is reduced to a 3-level hierarchy. And a final single value is obtained after two more iterations. In the following subsection we show how this procedure is applied to aggregate security criteria at level L3, since this branch of the taxonomy is more complete and thus suitable to better illustrate the benefits of the approach.

8.2.2. Test 2: Aggregation of L3 Metrics

For the Pre-Federation branch depicted in Figure 8.1, there are three criteria that have sub-criteria at hierarchy level 3 that must be aggregated: security, interoperability and knowledge. Here we show examples on the aggregation of security sub-criteria. To aggregate security characteristics into a single value that conveys risk information using our

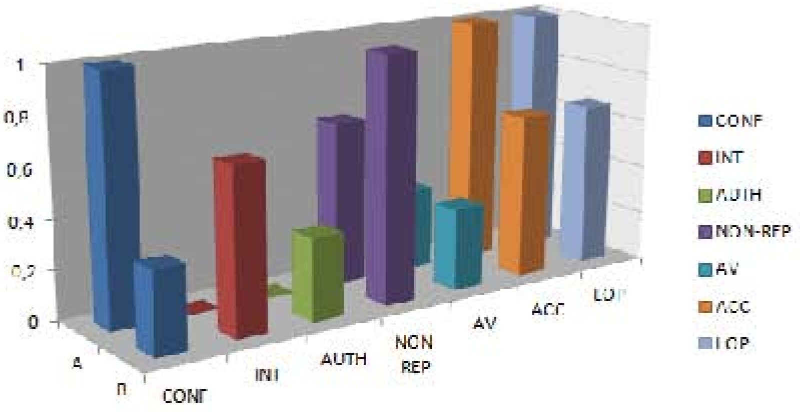


Figure 8.2: Representation of security assurance dimensions for two example entities

framework, the seven dimensions depicted in Figure 8.1 must be combined. Table 8.4 shows an example of two providers A and B with different security characteristics, also graphically depicted in Figure 8.2. The value for the aggregated *Security Assurance (SA)* is shown based on a reference vector associated to the evaluating party. The results that would be obtained by using the arithmetic mean are also shown to observe the differences.

SA Subcriteria	**SV_1^2 Provider A**	**SV_1^2 Provider B**	**RV_1^2**	**WV_1^2**
Confidentiality	3/3	1/3	1/3	1/8
Integrity	0	2/3	1/3	1/8
Authentication	0	1/3	1/3	1/8
Non-Repudiation	2/3	3/3	2/3	1/4
Availability	1/3	1/3	0	0
Accountability	3/3	2/3	2/3	1/8
Privacy	3/3	2/3	1/3	1/8
SA = Agg_1^2	0.66666667	0,458333333		
SA (%)	66.67%	45.83%		
ACI_1^2	0.71428571	1		
$CAgg_1^2$	0	0.458333333		
$CAgg_1^2$ (%)	0%	45.83%		
Mean	0.571428571	0.571428571		

Table 8.4: Aggregation of security assurance dimensions with minimum requirements constraints

Using the arithmetic mean operator, both entities provide the same security assurance, despite their security profiles are very different and only B fulfills the minimum requirements. This is graphically shown in Figure 8.3. As it can be seen, the usage of the weights allows to better rate the provider whose security criteria values are better according to the risk policy of the evaluating party. Furthermore, by means of the compliance index, the

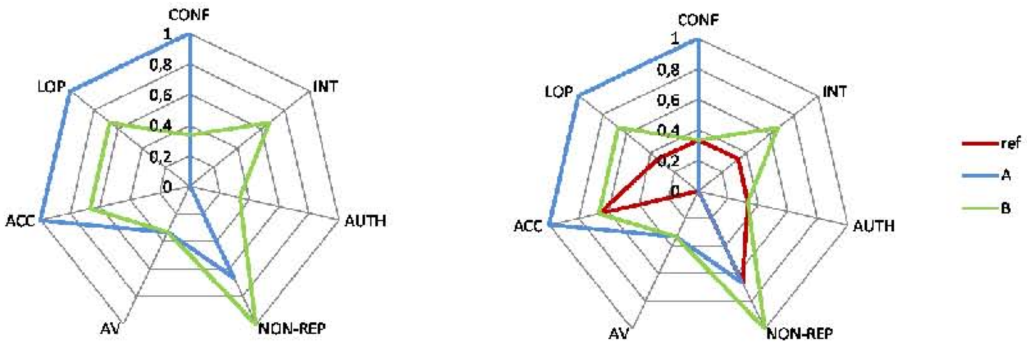


Figure 8.3: Score vectors for security sub-criteria of Provider A and Provider B (left), and their comparison with respect to a sample reference vector (right)

minimum risk requirements are directly embedded. Thus, the information conveyed by the aggregated value is twofold: on the one hand, it shows whether the minimum security risk requirements are satisfied; and on the other hand, when requirements are satisfied, it shows the security assurance level provided.

To give another example, let us assume that the evaluating entity risk policy says that the impact on an integrity attack is *High*, and the maximum integrity risk to be assumed is *Low*. This leads to a desired *Low* probability to incur in risks (see Table 8.3) and, consequently, to require a minimum confidentiality assurance equal to *High*. The rest of the security dimensions are not so relevant for the provider and no assurance is required as minimum. Its reference vector is then $RV_1^2 = [0\ 3/3\ 0\ 0\ 0\ 0\ 0]$. Graphically, it means that entities whose SV lines do not contain the RV line associated to the evaluating entity, do not satisfy its risk policies. For example, entities A and B described in the previous example do not satisfy the risk requirements, and this fact is reflected in their final $CAgg$ figure, which is 0 for both of them. However, an entity C with $SV_1^2 = [1/3\ 3/3\ 1/3\ 0\ 0\ 0\ 0]$, despite having a simpler security profile, it does satisfy the minimum risk requirements and so its ACI is equal to 1 and the $CAgg$ is different from 0. Figure 8.4 illustrates this example.

To give a last interesting example, for a reference vector $RV_1^2 = [1/3\ 3/3\ 1/3\ 1/3\ 1/3\ 1/3\ 1/3]$, a provider with score vector $SV_1^2 = [1\ 2/3\ 1\ 1\ 1\ 1\ 1]$ would have the same Agg value than a provider with score vector $SV_1^2 = [1/3\ 3/3\ 1/3\ 1/3\ 1/3\ 1/3\ 1/3]$. However, only the second provider satisfies the security requirements. The introduction of the ACI permits

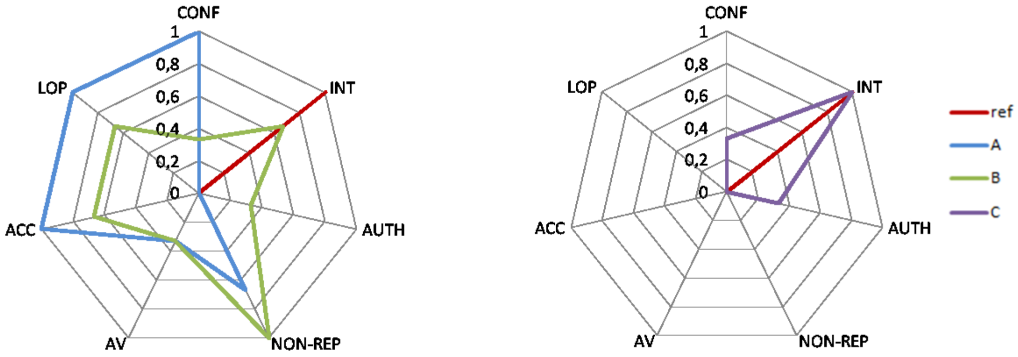


Figure 8.4: Score vectors for security sub-criteria of Providers A, B (left) and C (right) compared to a sample reference vector

to properly distinguish in this case, and to make a correct decision.

With this test, we can observe three important facts. Firstly, the aggregated values *Agg* depend on the reference vector, which allows us to model and include the subjectivity of the evaluating party. Secondly, the weights modulate the aggregation to better rate those dimensions that are more important for the evaluator. However, MAUT is a compensatory decision methodology because alternatives that are deficient with respect to one or more criteria can compensate by their good performance with respect to other criteria. In this sense, we see the importance of including the minimum requirements to overcome possible undesired choices as a side effect of this compensatory features of MAUT.

Having shown the aggregation procedure for the security branch, the method for the rest of the nodes is applied in exactly the same way. To conclude with this section, we give an example of the whole aggregation to calculate Pre-Federation Assurance.

8.2.3. Example: Assessing Pre-Federation Risk

In this example we will use real data about a federation named RedIRIS-SIR. This federation, operational since 2008, is composed by 102 IdPs and 200 SPs[1] operated by research and education institutions. Metadata documents and policies for joining the federation are publicly available at the RedIRIS website[2]. For the example, we have chosen two providers from the federation, namely 1) *University Carlos III de Madrid* as IdP, and 2) *Springer*

[1] Data from the federation survey in [TERENA, 2012], date 29February 2012

[2] http://www.rediris.es/sir/

online library as SP. Both are already federated after following a manual administrative process. Here we show how the risk could have been calculated before deciding on federation based on the available information. Thus, we show how the SP would evaluate the risk involved in establishing a federation with the IdP.

In order to apply out risk model, we start by summarizing the available information in the SP and IdP metadata and policies regarding the three dimensions in Pre-Federation. Based on this, we extract the metric values and construct the SP's reference vector and the IdP's score vector. Next, aggregation is performed to obtain the final risk value from the calculated vectors.

Information Gathering

The information obtained regarding the three dimensions of Pre-Federation is the folowing:

- **Security and Privacy.**

  The SP metadata declares support for XMLSignature with RSA signature algorithm and SHA digest algorithm. Also encryption is supported using the XMLEncryption standard with AES algorithm and 128 bits key size. The IdP metadata supports also RSA-SHA for signature, but no support for encryption is declared. However, despite the above mentioned cryptographic support, both indicate that assertion signing is not required.

  The mentioned data are related to confidentiality, integrity and authentication at message level. Information about the same dimensions at transport level is also found in RedIRIS policies, where the usage of TLSv1 or SSLv3 with a minimum of 128 bits key size and RC4 ciphersuite is recommended. Furthermore, the certificates used are RSA certificates with key size 2048. This affects the above security dimensions and also affects non-repudiation. With regard to the rest of security dimensions, few or no explicit information is provided. There are no data about availability. In the case of accountability it is required that the IdP collects appropriately timestamped logs containing at least requesting IP, user's NetID and opaque unique ID. Finally, in regard to privacy, the IdP must accept that the purpose of the federation is research (non-commercial), implement user consent and informed consent mechanisms (not a detailed policy for this), and adhere to the stipulations of the currently valid EU

Directive on Data Protection regarding processing of personal data when releasing attributes to 3rd countries.

- **Interoperability.**

  The SP and IdP metadata indicate support for a set of protocols, nameID formats, and bindings. Furthermore, a set of attributes that must be supported are detailed in the policy documents. All the mentioned data are related to technical interoperability. A summary of the main features is provided in Table 8.5.

Tech. Support	SP	IdP
Protocols	“urn:oasis:names:tc:SAML:2.0:protocol”	“urn:mace:shibboleth:1.0”, “urn:oasis:names:tc:SAML:1.1:protocol”, “urn:oasis:names:tc:SAML:2.0:protocol”
NameID formats	“urn:oasis:names:tc:SAML:2.0:nameid-format:transient”, “urn:oasis:names:tc:SAML:2.0:nameid-format:persistent”, “urn:oasis:names:tc:SAML:1.1:nameid-format:emailAddress”, “urn:oasis:names:tc:SAML:1.1:nameid-format:unspecified”, “urn:oasis:names:tc:SAML:1.1:nameid-format:WindowsDomainQualifiedName”, “urn:oasis:names:tc:SAML:2.0:nameid-format:kerberos”, “urn:oasis:names:tc:SAML:1.1:nameid-format:X509SubjectName”, “urn:mace:shibboleth:1.0:nameIdentifier”	“urn:mace:shibboleth:1.0:nameIdentifier”
Services/Bindings	SSO, SLO over HTTP-Redirect, HTTP-POST and SOAP	SSO, SLO over HTTP-Redirect
Attributes	ePTI, ePA, sHO, ePE and sPUC[3]	ePTI, ePA, sHO, ePE and sPUC

Table 8.5: Suported technical features by the SP run by Springer and the IdP run by University Carlos III

- **Knowledge.**

  In regard to knowledge, the available information declares that the CA that should be accepted as trusted for federation purpose is the Terena Certificate Service (TCS) [4].

Metrics Derivation

Based on the above data and a set of assumptions we will detail, the frameworks for the metrics can be refined and the values of these metrics obtained. The values for the

[4] https://www.terena.org/activities/tcs/

reference vector of the SP and the score vector of the IdP are shown in Table 8.6 according to the notation in Appendix B.

<table>
<tr><th>Metric</th><th>Vector Syntax</th><th>RV$_{IdP}$</th><th>SV$_{IdP}$</th></tr>
<tr><td>CCONF</td><td>[|| CONF$_{TL}$ ||, || CONF$_{ML}$ ||]</td><td>[1/3, 0]</td><td>[1, 0]</td></tr>
<tr><td>CINT</td><td>[|| INT$_{TL}$ ||, || INT$_{ML}$ ||]</td><td>[1/3, 0]</td><td>[2/3, 0]</td></tr>
<tr><td>CAUTH</td><td>[|| AUTH$_{TL}$ ||, || AUTH$_{ML}$ ||]</td><td>[1/3, 0]</td><td>[1/3, 0]</td></tr>
<tr><td>CSA</td><td>[CCONF, CINT, CAUTH, || NON-REP ||, || AV ||,|| ACC ||, ||LOP ||]</td><td>[1/3, 1/3, 1/3, 0, 0, 1/3,1/3]</td><td>[1, 2/3, 1/3, 0, 0, 1/3, 1/3]</td></tr>
<tr><td>CIA</td><td>[|| INTEROP$_{T}$||,|| INTEROP$_{O}$||,|| INTEROP$_{L}$||]</td><td>[1/3, 0, 0]</td><td>[1/3, 0, 0]</td></tr>
<tr><td>CKNOW</td><td>[|| KNOW$_{D}$||,|| KNOW$_{I}$ ||]</td><td>[1, 0]</td><td>[1, 0]</td></tr>
<tr><td>PreFedA</td><td>[CSA, CIA, CKNOW]</td><td>[1/3, 1/3, 1]</td><td>[8/15, 1/3, 1]</td></tr>
</table>

Table 8.6: Score Vectors for Aggregated Pre-Federation Metrics

For the calculation of the values in Table 8.6, we have considered that:

- The minimum ciphersuite required is TLS_RSA_WITH_RC4_128_SHA. However, after testing the support by the IdP server, we saw that it can provide as highest strength ciphersuite DHE-RSA-AES256-SHA. According to this and to the rest of the security information, the values for the security metrics were obtained following the methodology in Chapter 5.

- We assume an Accountability Assurance Framework for the SP that requires as minimum log information: timestamp, requesting IP, user's NetID and opaque unique ID. Other elements such as time for caching the logged data would be optional elements in the framework. We assume that the IdP satisfies the minimum requirements in LOP and does not provide any additional feature.

- We assume a LOP Assurance Framework for the SP that is composed by elements *purpose*, *user informed consent* and *privacy directives*. The purpose can take the values "commercial", "non-commercial". The user informed consent can take values "non required", "at IdP discretion", "uApprove"[5]. The privacy directives can take the value "EU Directive on Data Protection". According to the information gathered, the values "commercial", at "IdP discretion" and "EU Directive on Data Protection" are the minimum required, while the rest would be optional. We assume that the IdP satisfies the minimum requirements in LOP and does not provide any additional feature.

[5] uApprove is an implementation of a user consent module for Shibboleth IdPs (`http://www.switch.ch/aai/support/tools/uApprove.html`)

- We assume a Technical Interoperability Assurance Framework for the SP composed by the elements in table 8.5: *protocols*, *nameID formats*, *services/bindings* and *attributes*. The possible values for each element are those shown in the SP column. In the case of the protocol element, "urn:oasis:names:tc:SAML:2.0:protocol" is required. In the case of nameID formats, it is required that at least one of the formats is supported and the rest are optional (the higher the coverage percentage, the higher the interoperability value). In the case of services/bindings, we assume that SSO over HTTP-Redirect is the minimum required.

  Furthermore, regarding operational interoperability, no data are available to define a framework so we assume that there are no requirements. Also, for the legal interoperability, since no other laws apart from the privacy regulations (covered in the LOP metric) are mentioned, we assume that there are no further legal requirements.

- We assume that the entities need to have and have a valid TCS root certificate that is trusted. So direct knowledge must exist, but no any additional indirect information is required.

Risk Aggregation

Applying the aggregation model to the values in Table 8.6, the final Pre-Federation assurance value is 0,773333333 (or an assurance of 77.3%) . It reflects that all the minimum assurances required by the SP in the different dimensions of the hierarchy are covered by the IdP (otherwise the value would be 0). And gives an idea of the provided assurance in a quantitative scale from 0 to 1.

This example together with the rest of the performed tests prove that our model works according to the design goals. But to validate results in the context of risk assessment, apart from testing that the model is correct in regard to the design requirements, researchers have proposed other three main approaches [Fenz and Ekelhart, 2011]: 1) using experts, 2) using alternate decision processes, and 3) using statistical evidence. The first approach, nominating technology experts, involves asking them to compare the findings of an applied approach with the results they would expect based on their intuition. The second approach consists of running at least one alternate decision process on the exact same problem. Finally, the statistical evidence approach involves analyzing internal reports and incidence related data and test the model against them to see if the risk-based decisions are correct.

Approaches 2) and 3) are more complicated since usually no other tool for decision making is available tailored to the scenario under analysis and, in the case of statics, organizations often refrain from making these data publicly available since the publication of threats incidence (such as data loss and fraud) may lead to a loss of reputation. This is exactly the case for FIM, i.e., no other risk-based decision tools exist and threat data are not publicly available.

The first approach, though more subjective than the others, is easier to accomplish and will be our next step in validation. The results so far, i.e., the definition of the taxonomy, the initial aggregation model the and metrics, were already disseminated in publications [Arias, 2011][6] and [Arias et al., 2012b].

8.3. Trust Model Validation

This section is dedicated to the validation of the trust model proposed in Chapter 6. More specifically, we want to prove the benefits of including reputation data in FIM networks. This aspect is complex to test in a real world scenario since it is would imply the deployment of a complex infrastructure with a high number of providers, which is not a trivial task. Thus, we have opted for a simulation-based validation using OMNeT++ [OMNeT++, 2012][7], which is an open source C++ simulation framework widely used in the scientific community.

This section describes the simulation tests performed with OMNeT++, showing relevant aspects of the system behavior. The results obtained have been analyzed in order to validate the feasibility of the proposed framework. Finally, we summarize the results and give recommendations on how to implement the protocol depending on the features of the underlying FIM network.

[6] This work received the *Best PhD Forum Contribution Award* at PERCOM conference 2011 (http://www.percom.org/2011/)

[7] The OMNeT++ model for testing the reputation protocol over FIM networks was developed during a research stay at NEC Laboratories Europe in Heidelberg

8.3.1. Goals and Metrics

As an initial goal, we wanted to set up experiments for the reputation protocol in order to find a balance between a good accuracy and a good overhead. To this end, we made several studies to investigate the scalability of the solution and the impact of introducing different fractions of malicious nodes in the network. For future work our goal is also to learn how other parameters (such as thresholds, trust formulas, contexts, etc.) should be optimally tuned.

We consider the following metrics in order to analyze and evaluate the IdMRep protocol:

- The **message overhead**. We look at the overhead caused by the extra messages issued when using the reputation protocol. The message overhead for a node i (MO_i) in a FIM network is calculated as shown in equation 8.4:

$$MO_i = \frac{RepRequest_i + RepResp_i}{IntendedTransactions_i} \tag{8.4}$$

  where $RepRequest_i$ and $RepResp_i$ refer to the number of `ReputationRequest` and `ReputationResponse` messages sent by node i, and $IntendedTransactions_i$ is the number of transactions initiated by the node. We use this metric to determine how much extra overhead IdMRep adds regarding to the regular overhead in a FIM network, i.e., the cost of introducing a new protocol.

- The **accuracy** (or success rate), which reflects the percentage of successful transactions. We calculate the success rate for a node i (SR_i) as the number of interactions with good entities plus the avoided interactions with malicious entities over the total number of transactions performed by the node, as shown in equation 8.5:

$$SR_i = \frac{TransGood_i + AvoidedTransMal_i}{TotalTransactions_i} \tag{8.5}$$

  where $TransGood_i$ is the number of transactions accepted by node i coming from a good entity; $AvoidedTransMal_i$ is the number transactions rejected by node i and coming from malicious entities; and $TotalTransactions_i$ is the total number of transactions in which node i can participate during the running of a simulation (i.e., for a SP, the number of `ReputationRequest` messages sent; and for an IdP, the number of `ReputationRequest` messages received). We use this metric in order to reflect the number of correct decisions made, i.e., how accurate the protocol is.

8.3.2. Simulation Setup

For the performance analysis of the reputation protocol, the metrics defined in Section 8.3.1 have been measured in various scenarios. First of all, we differentiate between two kinds of FIM networks:

A) networks where there are relationships only between SPs and IdPs; and

B) networks with SP-IdP and IdP-IdP relationships. Both types of network graphs are represented in Figure 8.5 and Figure 8.6, respectively.

A relationship between providers means that they can transact with each other following a federation protocol. We chose these two network cases because, according to FIM protocols, there can be relationships only between IdPs and SPs (IdPs send user information to SPs in order for the users to gain access to services), but also there can exist relationships between two IdPs (e.g., an IdP may delegate some identity tasks to another IdP). As we will see later, the simulation results show that network type B improves the connectivity and so the obtained accuracy is better.

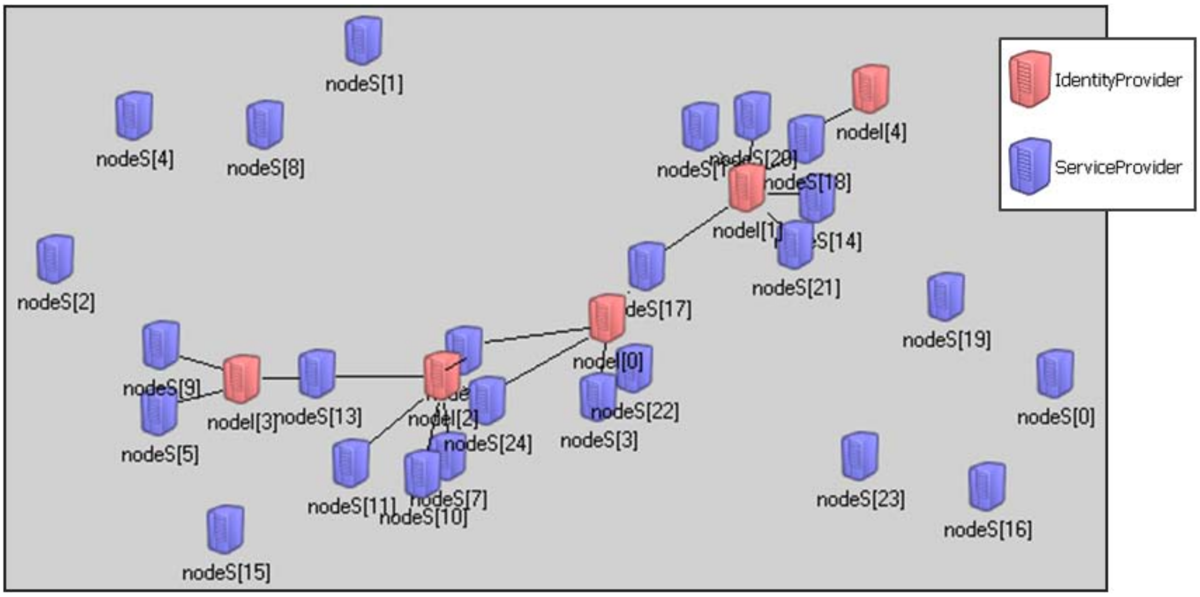


Figure 8.5: FIM network of type A with 25 SPs and 5 IdPs.

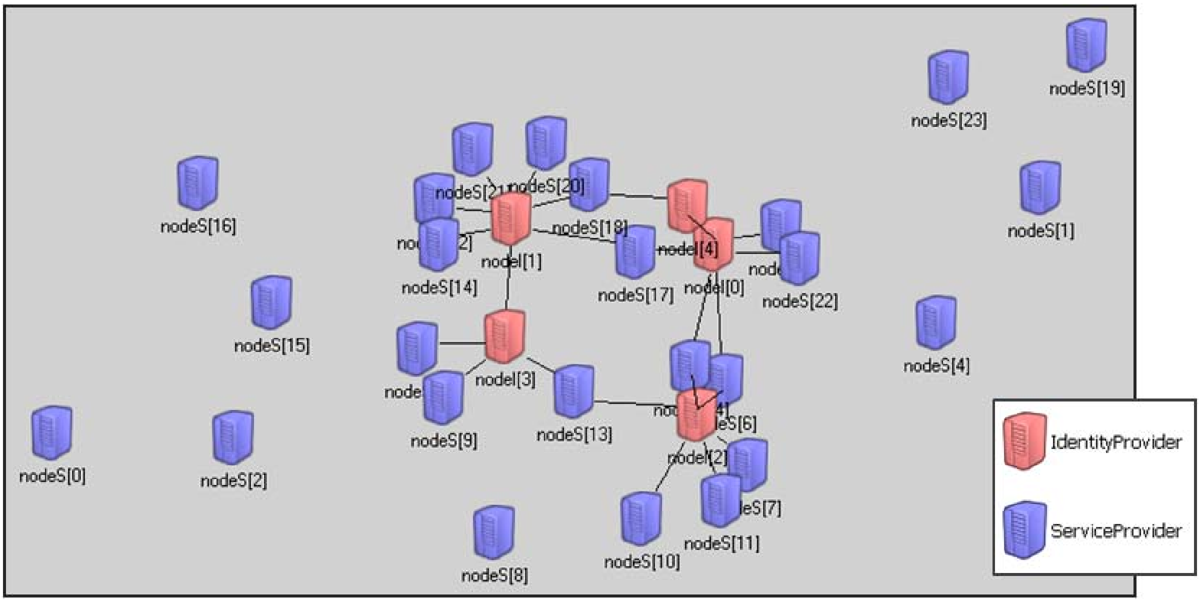


Figure 8.6: FIM network of type B with 25 SPs and 5 IdPs.

Besides, Table 8.7 compiles all the possible parameters that can be configured regarding network model, trust and reputation model, entity behavior model, IdMRep forwarding model and regarding the simulation itself. In order to simplify the experiments we use the formulas 6.1 and 6.2 presented in Chapter 6 for trust and reputation aggregation assuming that:

- T_{Auth} is always equal to 1. We assume authentication is always correct and evaluate only the behavioral part.
- The evolution function f_{evol} is a lineal aggregation function in the form shown in equation 8.6:

$$f_{evol}(i) = \alpha \cdot trustworthiness(i) + (1 - \alpha) \cdot sat(i)) \tag{8.6}$$

  Where α is an adjustable parameter used for tuning the importance of recent and older transactions. The same evolution function is used to compute the reputation using as input the set of received votes.

The configurable parameters regarding trust and reputation shown in Table 8.7 are related to the mentioned formulas, but any other evolution formula could be used. Furthermore, we consider for this set of initial experiments that only SPs can act maliciously, since this is the most probable scenario .

FIM Network Model	#IdPS	Number of Identity Providers
	#SPs	Number of Service Providers
	ConnectivitySI	Connectivity Degree between SPs and IdPs, defining the preexisting relationships
	ConnectivityII	Connectivity Degree between IdPs and IdPs, it defining the preexisting relationships
	Network Type	Connections between SPs-IdPs (A) or Connections between SPs-IdPs and IdPs-IdPs (B)
Trust and Reputation Model	d	Trust disposition, i.e., initial default trust value assigned to an entity when no reputation data about it are available
	α	Parameter in the trust model to adjust the importance of new transactions and old transactions
	MALTHRESHOLD	Trust threshold for decision making (i.e., if local trust for entity i is greater or equal than MALTHRESHOLD, then a transaction can be initiated/accepted)
Entity Behavior Model	MalRate	Fraction of malicious entities in the Network
IdMRep Forwarding Model	TTL	Time To Live for reputation Requests
	N	Number of Nodes in the DTL to forward reputation requests
Simulation	#repetitions	Number of repetitions per experiment

Table 8.7: Configurable Parameters

The operation for each simulation running implies that nodes randomly chose another node in the network to transact with and, according to the logic explained in Chapter 6, the IdMRep protocol is executed if there is no entry in the DTL for the chosen node. After gathering reputation opinions, the initial node makes a decision whether to transact or not. This behavior is constantly repeated through the duration of the experiment. The simulation time is chosen to be long enough to potentially cover a huge number of interactions.

We also introduce different fractions of malicious entities in the system in order to analyze the impact on the success rate. Each experiment is repeated 5 times with varying random seeds. The seed influences the order in the sequence of initiated transactions. Whenever the confidence intervals are shown in the plots, the confidence level of these intervals is 95%. To summarize, Tables 8.8 and 8.9 show the list of fixed parameters and the list of variable parameters chosen for the configurations used in the performed experiments, respectively.

Parameter	Value
SPs	25
ConnectivitySI	0.2
ConnectivityII	0.075
MALTHRESHOLD	0.5
α	0.5
N	Size of DTL
#repetitions	5

Table 8.8: Fixed parameters used in the simulations.

Parameter	Value
#IdPs	5, 10, 15, 20, 25
Network Type	A,B
d	0.5, 0
MalRate	0.1, 0.2, 0.3, 0.4, 0.5
TTL	1, 2, 3, 4, 5, 6

Table 8.9: Configurable parameters used in the simulations.

Basically, we look at the overhead and accuracy under different choices of TTL and considering the impact of the fraction of malicious nodes (MalRate) and of the number of IdPs (#IdPs) in the network. We do that for the two types of network graphs and also using two possible values for the initial trust (d) to be placed in unknown entities.

8.3.3. Simulation Results

Initial Results under “Nice” Conditions

Before performing the set of experiments described in the previous section, we first analyze the protocol and take measurements supposing “nice” conditions in the simulated environment, that is, without adding malicious nodes.

For this analysis we decide to choose the parameters in Table 8.10. The main goal is to observe the behavior of the protocol when varying the TTL in terms of overhead and percentage of successful transactions. In this case we use refer to the accuracy metric as the “transaction rate” because when there are no malicious entities, all the completed transactions will be successful. Thus, the transaction rate reflects the percentage of completed transactions over the total of intended ones.

We performed two opposed experiments: (a) with entities willing to cooperate even if no reputation data are found about other unknown providers (i.e., d = 0.5); and (b) with

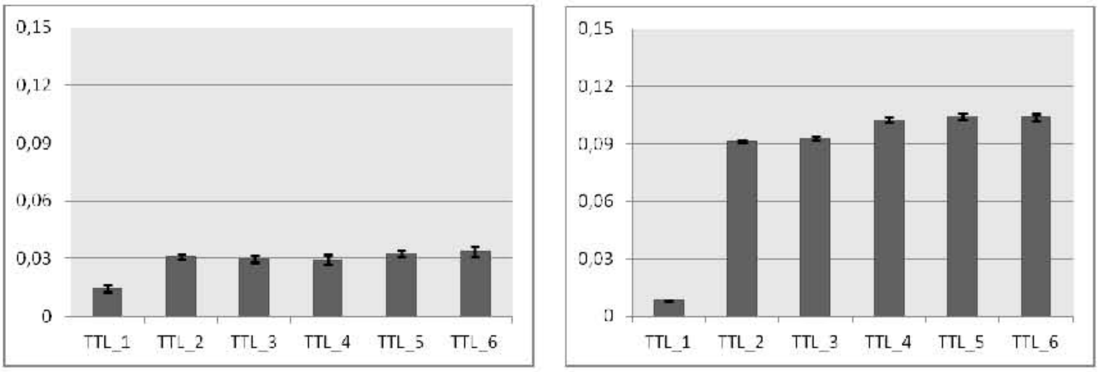
Parameter	Value
#SPs	25
#IdPs	5
#Network Type	B
ConnectivitySI	0.2
ConnectivityII	0.075
MALTHRESHOLD	0.5
d	0.5, 1
α	0.5
N	Size of DTL
Malrate	0
#repetitions	5

Table 8.10: Fixed parameters used for the simulations supposing nice conditions.

entities that only transact with other unknown entities if reputation data are available (i.e., d= 0). The aim of these experiments is to show the benefits of IdMRep with respect to the current FIM frameworks. Thus, we previously identified that there are two FIM models nowadays: the *"accept-all-comers"* model, which means that entities always trust every other unknown entity; and the rigid trust model, which means that no other entity will be considered trusted, except those that have been manually pre-configured . Hence, with experiment (a) we test the benefits and disadvantages of applying IdMRep to the *"accept-all-comers"* model. Similarly, with experiment (b) we test the benefits and disadvantages of introducing the reputation protocol in a rigid FIM model.

The results for experiment (a) are shown by the graphs in Figures 8.7 and 8.8. The average transaction rate is 100% since entities always trust. Thus, in a nice environment may seem useless to introduce reputation, but the results in the next section will show the benefits of the approach when there are malicious entities in the system. Also, in a "nice" environment, different good reputation values may help to decide on giving or denying different kinds of transactions. From the point of view of the overhead, we can see that this is not high: once the entities are recognized as trustworthy and inserted in the DTLs, the overhead remains low.

On the other hand, the results for experiment (b) are shown by the graphs in Figures 8.9 and 8.10. The average transaction rate in this case is not 100%, since now entities do not trust other entities if there is no reputation available. Instead, the transaction rate value is around 60% for all TTL cases, which drastically improves the case of using a rigid trust model where the percentage of transactions with unknown entities is 0%. As a negative counterpart, the average overhead for SPs is a bit higher than the obtained in the case of using a trust disposition factor = 0.5. It is to note that in any case, the overhead is



(a) Average overhead for SPs

(b) Average Overhead for the IdPs

Figure 8.7: Average overhead for SPs(a), and for IdPs(b) when varying the TTL parameter, d = 0.5

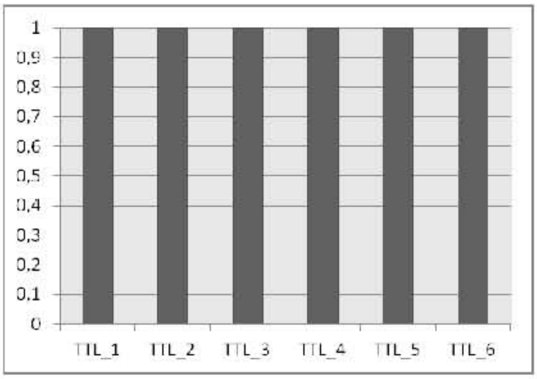
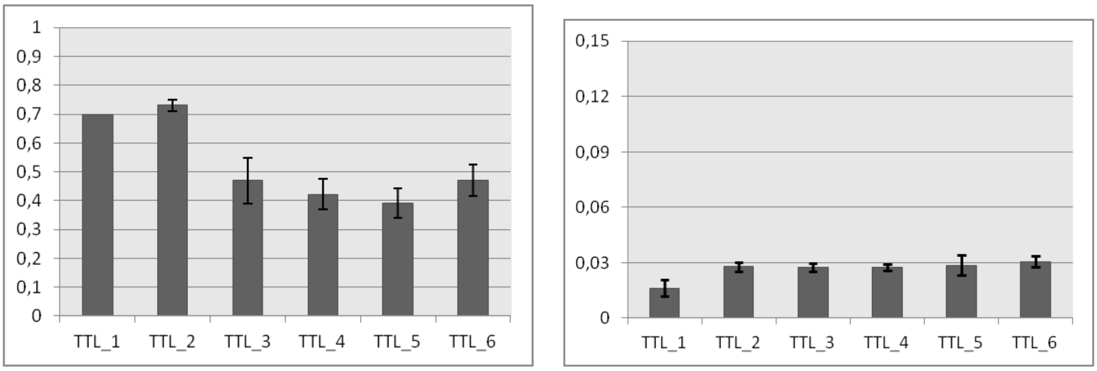


Figure 8.8: Average transaction rate when varying the TTL parameter, d = 0.5

reasonably low to positively consider the introduction of the protocol in FIM networks. Furthermore, the TTL value does not influence significantly the overhead. This behavior will be also observed later in the rest of the experiments.

To conclude, in case (a) the transaction rate is the same as if we do not apply the protocol but the reputation will help to identify malicious entities, as we will see in the experiments, and also to provide entities with different granularity levels regarding the trustworthiness that can be placed in another entity. It can help in allowing/denying transactions depending on the trustworthiness levels.

Hence, the proposed protocol allows us to achieve a tradeoff between the *"trust-all-comers"* and the rigid trust model, imposing a reasonably good overhead. After analyzing and proving the usefulness of the proposed framework under "nice" conditions, we present a



(a) Average Overhead for the SPs (b) Average Overhead for the IdPs

Figure 8.9: Average overhead for SPs(a), and for IdPs(b) when varying the TTL parameter, d = 1

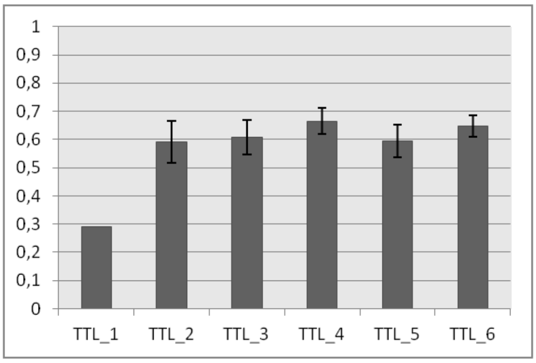


Figure 8.10: Average transaction rate when varying the TTL parameter, d = 1

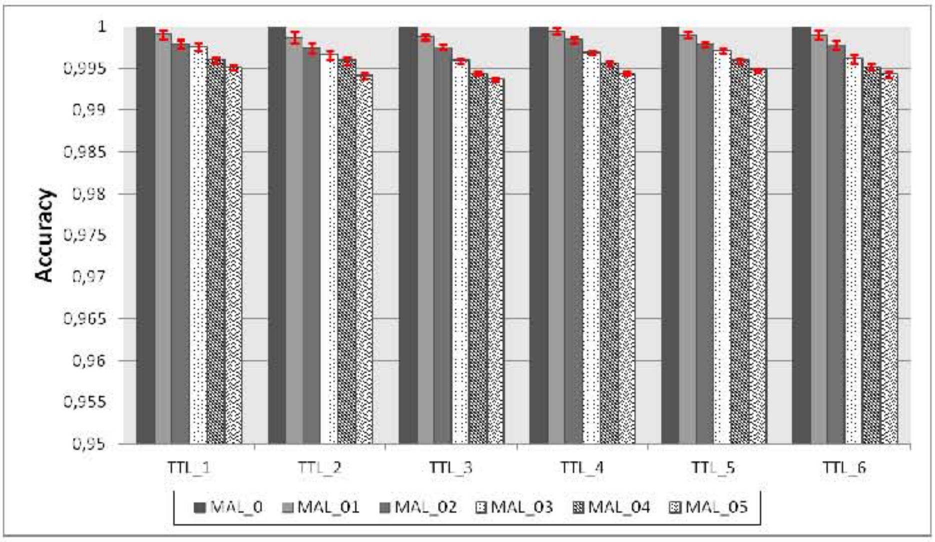
set of tests designed to study the influence of malicious nodes in the system.

Results under Impact of Malicious Entities

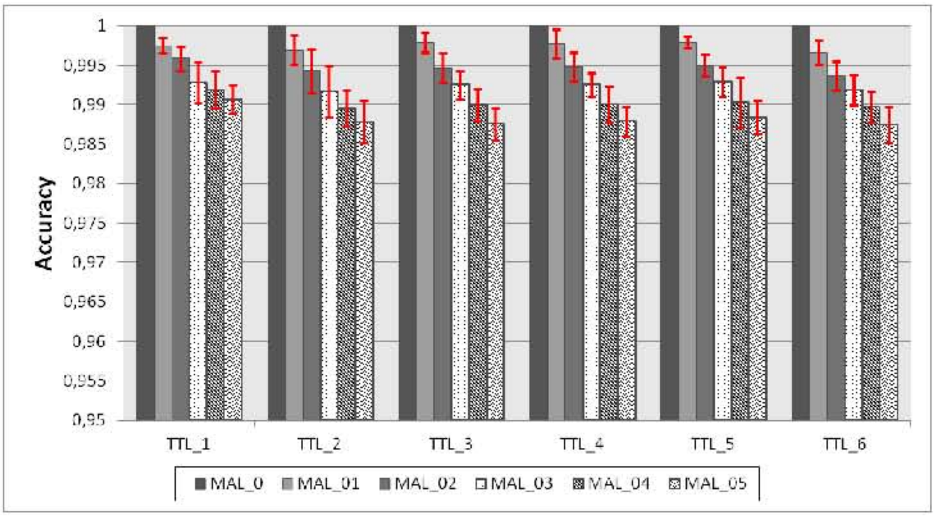
Aim to study the impact of certain parameter variations, namely the TTL value, the percentage of malicious entities, the number of IdPs and the type of network.

In Figures 8.11 and 8.12 we show the accuracy obtained in a network of type A, where the *d* value assigned to unknown entities is 0.5. The number of IdPs is increased from an initial value of 5 to a final value of 20 according to Table 3. Besides, the results are presented for TTL values between 1 and 6, and percentage of malicious nodes ranging from 0% to 50%.

Having 5 IdPs, the obtained accuracy is always over 99% and it decreases with the percent-



(a) Accuracy for increasing percentage of malicious nodes varying the TTL, Network type A, d=0.5, #IdPS=5

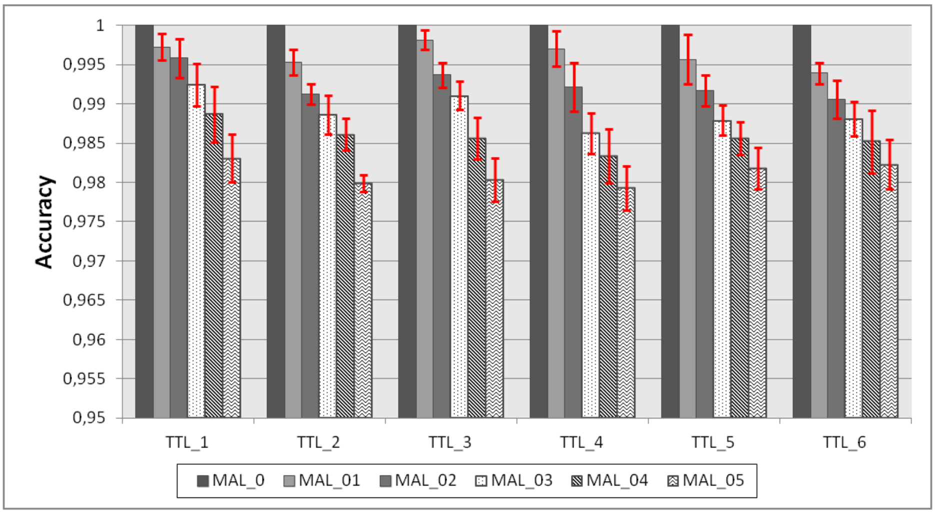


(b) Accuracy for increasing percentage of malicious nodes varying the TTL, Network type A, d=0.5, #IdPS=10

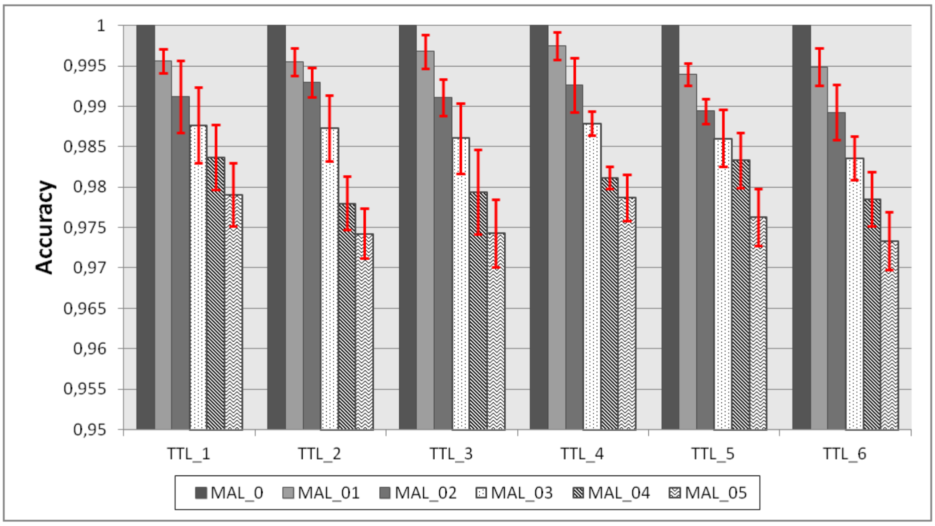
Figure 8.11: Accuracy measured in Network type A with parameter $d = 0.5$ varying TTL, MalRate and #IdPs= 5,10

age of malicious entities. As it can be observed, increasing the TTL value does not have a significant impact on the accuracy. The same tendency regarding malicious behavior and TTL variations is observed in the rest of the graphs for networks with 10, 15 and 20 IdPs. On the other hand, the accuracy gets lower when the number of IdPs increases: around 98% for 10 IdPS, 97.5% for 15 IdPs and 97% when the number of IdPs reaches 20.

In the following, we do not show the graphs for the rest of the simulation cases but discuss



(a) Accuracy for increasing percentage of malicious nodes varying the TTL, Network type A, d=0.5, #IdPS=15



(b) Accuracy for increasing percentage of malicious nodes varying the TTL, Network type A, d=0.5, #IdPS=20

Figure 8.12: Accuracy measured in Network type A with parameter $d = 0.5$ varying TTL, MalRate and #IdPs=15,20

the obtained results instead. Thus, the next test consisted of applying the same configuration of simulation parameters as in Figure 7, but using a d value equal to 0, i.e., providers will not cooperate with unknown entities. The results obtained in this set of experiments are very similar to the ones obtained when using $d = 0.5$: the accuracy tendencies regarding TTL variation, increasing percentage of malicious nodes and increasing number of IdPs are the same. The main difference is that in this case we have a slightly more accurate system, but with the negative counterpart that the number of transactions is reduced (as shown in Figure 8.10 for the “nice” environment). On the other hand, we saw that there is no difference in the overhead regarding the percentage of malicious entities. From the outcomes, we have also observed that the overhead is always lower for IdPs than for SPs. However, as the number of IdPs increases, the average overhead for SPs decreases and so both overheads tend to be similar.

The overhead increases with the TTL but not significantly. Selecting a TTL equal to 1 in networks of type A means that entities will never get reputation data, so there is no sense in choosing this value. In the case of networks of type B, entities can get reputation data in one hop and the overhead is dramatically lower. However, selecting a greater TTL increases the chances of cooperation and better knowledge of the network.

The difference between a network with nodes tending to trust versus a network with conservative nodes (i.e., using d equal to 0.5 or 0, respectively) is that the overhead for the latter case increases. This is due to the fact that many times IdMRep messages are sent but the interaction is finally not completed because no reputation data is obtained.

The behavior in terms of accuracy and overhead regarding the variable parameters is very similar in both network types. We decided to choose a connectivity degree of 0.075 between IdPs because this degree is usually lower than the connectivity between SPs and IdPs. Thus, although the results show little changes in networks of type A and B, more experiments with different degrees of connectivity are required for a definitive conclusion. As future work, we plan to make more experiments to better analyze the impact of applying the protocol in different types of networks.

Besides, the obtained accuracy is very high in any case. But it is to consider that these results are conditioned to a static behavior of malicious nodes: once the entities are recognized as trustworthy and inserted in the DTLs, the overhead remains low. On the other hand, using a d value equal to 0.5 increases the number of transactions and improves cooperation with unknown entities, still maintaining a good accuracy.

In general, it was expected that increasing TTLs lead to increasing accuracy and overhead. However, the variation of the results with this parameter is small. In fact, the overhead value depends more on the disposition of the entities to trust. For example, if entities are not willing to trust unknown entities ($d = 0$), the knowledge of the network is based only on the pre-existing relationships, so the procedure is slower and more `ReputationRequests` are sent that do not lead to a transaction. So, regarding TTL selection, we conclude that any value greater than 1 provides similar results in accuracy and overhead, and this is due to the special connectivity features of FIM networks. The recommendation is to chose high values for the TTL because the overhead does not increase significantly and they lead to a better knowledge of the network and higher number of transactions. To conclude with the discussion, the overhead imposed by the protocol is assumable and makes the solution feasible. Furthermore, the obtained accuracy is very good (over 94% in the worst case), since once the system learns, the transactions always succeed.

However we should analyze scenarios where the malicious pattern is not static and study its impact on the accuracy. It would be also useful to have real data about FIM deployments in order to better set the parameters of the simulations so they can reflect real world configurations, but there is not available public data in this regard.

Summarizing, we presented a set of initial experiments that constitutes a foundation to continue building and improving dynamic cooperation in FIM as described in this this thesis. Thus, our further research lines in this regard, which will be furtherly explained in Chapter 9, aim to cover more complex scenarios in order to enhance and adapt the reputation protocol.

8.4. Architecture Validation

According to [Ambler and Lines, 2012], the best way to validate an architecture model is through implementation. The questions to be answered in this phase are:

1. Does the architecture work?

2. Is the approach feasible?

With the aim to answer the above questions, we have developed an architectural prototype that partially implements selected modules of the whole architecture proposed in Chapter 4.

8.4.1. Conceptual Description: Dynamic SSO Using Reputation Data: a *Proof-of-concept*

As a first approach we have developed a proof-of-concept application to show the viability of our dynamic federation architecture proposal. For this purpose, we work with the simplest SAML-based SSO scenario: a user, an IdP and a SP. The providers are unknown to each other, i.e., no formal contracts or trust pre-configuration exists. In this situation, current providers either avoid interaction (rigid trust) or interact by default (trust-all-comers).

In the specific case of the identity management infrastructure used for the experiments, the user must introduce the URI where the metadata document of her IdP is located. This must be done when the providers are unknown. After that, metadata are stored and all the certificates contained in it are considered trustworthy for validating subsequent SAML messages. The main limitation regarding trust issues is that no trust model is defined since the SP will interact with any provider introduced by the user.

Thus, to enable dynamic and secure interaction we have added the new trust logic that modifies the usual operation diagram. It has been done through the implementation of reduced versions of all the designed architecture modules, except from the *Risk Manager*.

The original work-flow and the modified one are depicted in Figure 8.13. As it can be seen, the explained *proof-of-concept* application serves to prove that the core ideas of our proposal are workable. In the following, we will explain the implementation details.

8.4.2. Implementation

In order to evaluate our proposal, we have deployed our own identity management infrastructure (see Figure 8.14). For the IdP, we are using Authentic [Authentic, 2007], which is a Liberty-enabled identity provider based on the lasso library [lasso, 2010] that also supports SAML 2.0 specifications. Authentic is a Quixote application, which is a framework for developing web applications in Python. We have also integrated Authentic with OpenLDAP [OpenLDAP, 2013] to manage users' accounts. The required digital certificates were generated by means of a certification authority deployed using the OpenCA [OpenCA Research Labs, 2013] software. For deploying SPs we have chosen ZXID [SymLabs, 2012],

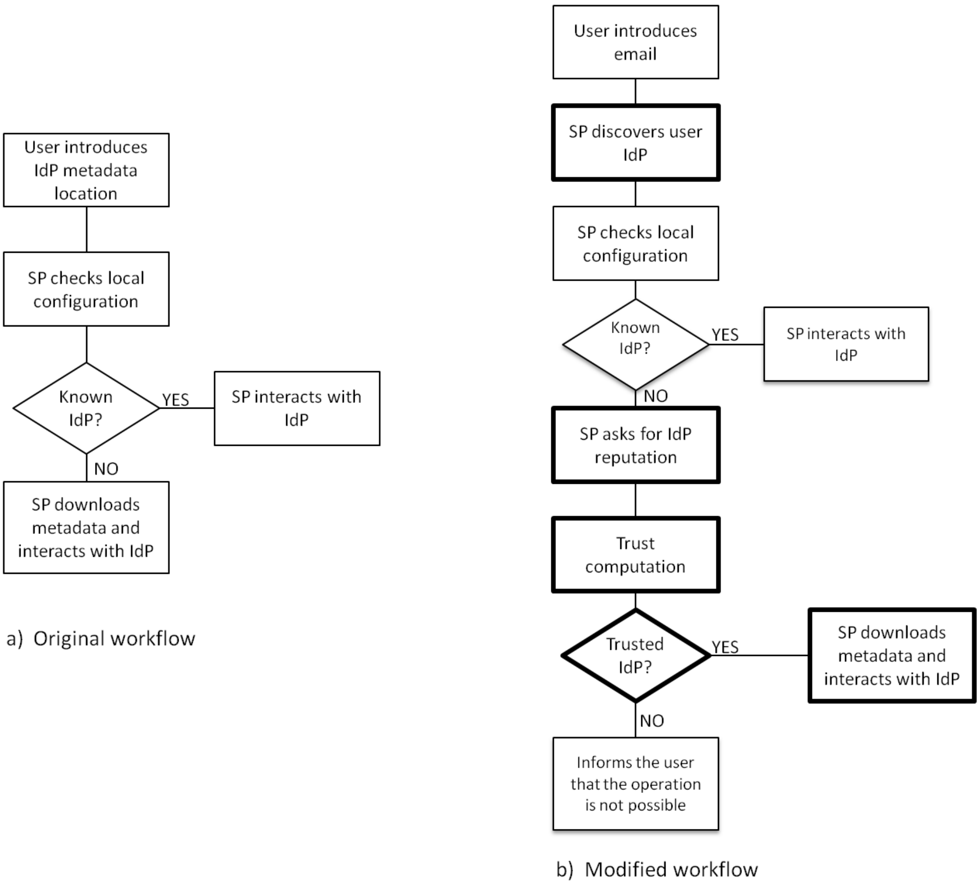

User introduces IdP metadata location
SP checks local configuration
Known IdP?
YES
SP interacts with IdP
NO
SP downloads metadata and interacts with IdP
a) Original workflow
User introduces email
SP discovers user IdP
SP checks local configuration
Known IdP?
YES
SP interacts with IdP
NO
SP asks for IdP reputation
Trust computation
Trusted IdP?
YES
SP downloads metadata and interacts with IdP
NO
Informs the user that the operation is not possible
b) Modified workflow

Figure 8.13: a) Original work-flow followed by the SP used in the evaluation; and b) modified work-flow after our extension is implemented in the SP.

a light open C library that implements the full SAML 2.0 stack. Both IdP and SP selected implementations are open source and use OpenSSL [OpenSSL, 2013] as underlying cryptographic library and Apache2 [Apache, 2012] as the web server.

Based on the described infrastructure we have introduced the modifications to implement the proposed architecture. For this purpose, we have developed the *Trust Manager* component, including the functionality to ask for reputation data. We have implemented the `ReputationRequester` and the `ReputationResponder` functions, which are software processes that can be added to SAML entities. As the names suggest, the requester is in charge of asking for reputation data regarding a provider and the responder is the responsible of answering to reputation requests. Both the format of the *Reputation Assertions*

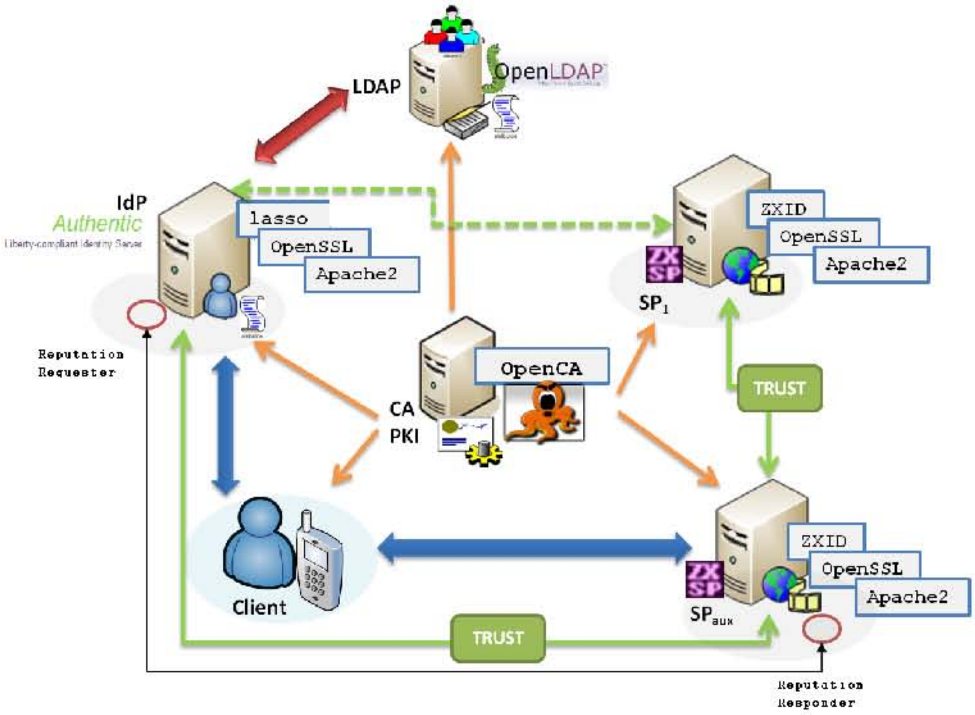


Figure 8.14: Implemented *proof-of-concept*

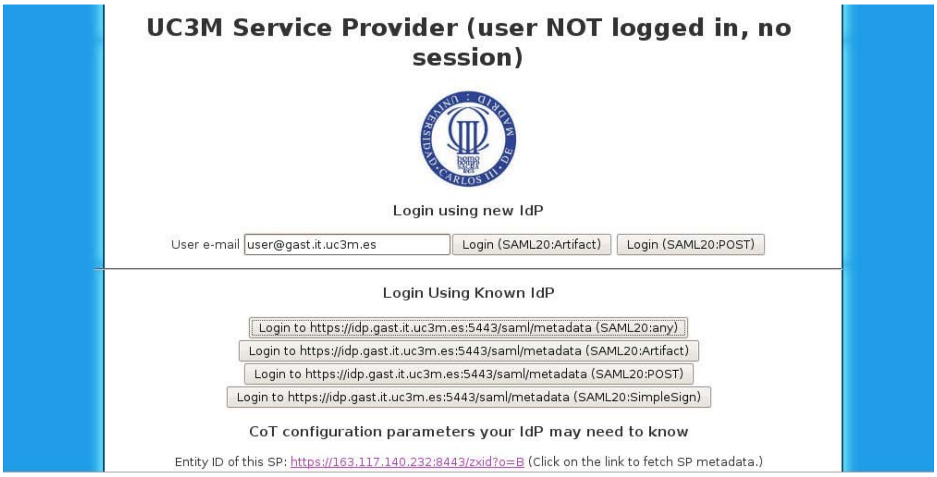
and the exchange protocol, follow the description provided in Chapter 6. The concept of DTL has been also implemented as a hash structure containing the known entities and the associated trust information. For the specific use-case we tested, despite the central scenario is SP-IdP-User, we introduce another service provider (SP$_{aux}$) that has a relationship with the IdP and thus possesses information about its reputation. The relationships stored on the DTL of each provider are shown in Table 8.11

DTL	Known Providers
DTL(SP)	SP$_{aux}$
DTL(SP$_{aux}$)	SP, IdP
DTL(IdP)	SP$_{aux}$

Table 8.11: Information contained in the DTLs of the providers involved in the implementation test

So, the SP has been modified to use the new trust functionality in the decision-making process in order to dynamically determine whether to cooperate or not. Then, we tested how changing the reputation value associated to the IdP in the DTL accessed by the `ReputationResponder` function in SP$_{aux}$, or adjusting the threshold reputation value in the SP, the provider makes different decisions to cooperate, always in a dynamic fashion. Communication was tested over different protocol bindings, such as HTTP-Redirect or POST.

It is to mention that the application also simulates the IdP and metadata discovery processes depicted in diagram of Figure 8.13. We use a function to derive the IdP name from the user email, but any other discovery service or application could be used since the SAML specification sets themselves are neutral with respect to any particular IdP discovery mechanism. A screenshot of the user interface is shown in Figure 8.15.



UC3M Service Provider (user NOT logged in, no session)

Login using new IdP

User e-mail user@gast.it.uc3m.es

Login (SAML20:Artifact)
Login (SAML20:POST)

Login Using Known IdP

Login to https://idp.gast.it.uc3m.es:5443/saml/metadata (SAML20:any)
Login to https://idp.gast.it.uc3m.es:5443/saml/metadata (SAML20:Artifact)
Login to https://idp.gast.it.uc3m.es:5443/saml/metadata (SAML20:POST)
Login to https://idp.gast.it.uc3m.es:5443/saml/metadata (SAML20:SimpleSign)

CoT configuration parameters your IdP may need to know

Entity ID of this SP: https://163.117.140.232:8443/zxid?o=B (Click on the link to fetch SP metadata.)

Figure 8.15: User interface of the *proof-of-concept* implementation

8.4.3. Further Validation and Lessons Learned

Both the *proof-of-concept* prototype, as well as the architectural design and research ideas on which the prototype is based, were contributed to the national R&D project "España Virtual"[8]. España Virtual is a CENIT[9] project funded by CDTI (Centre for the Development of Industrial Technology), Spanish Government.

The main goal of the project, that lasted between 2008 and 2012, was to establish a bridge between geography and the Internet through the definition of an architecture, protocols and standards of Internet geography, with a special focus on processing data, 3D visu-

[8] http://www.españavirtual.org/

[9] The National Strategic Consortia for Technical Research (CENIT) programme is promoted by the Spanish Government, within the context of its programme to foster long-term public-private cooperation in the field of R&D, with the objective of promoting and developing industrial research

alization, virtual worlds and interaction between users. An specific working package for "*Security and Identity Management*", where our ideas were developed [España Virtual, 2010a], was included to give a flexible and secure support to the envisioned rich ecosystem of services.

In addition, the part of the architecture including reputation, together with the simulation results presented in Section 8.3, were contributed in deliverables of an internal business project centered in IdM at NEC Laboratories Europe.

Once we tested the architecture depicted in 8.14, we defined and developed new use-cases to be integrated in the scenarios of the project. More specifically, besides the SPs and IdPs located in the domain of University Carlos III and in one of the participant companies, we offered the possibility of introducing external third parties (e.g., Facebook, Twitter) as IdPs. We successfully tested the dynamic federation with these 3rd parties, and the subsequent delegation of user authentication to them.

Furthermore, not only a browser-based used interface was implemented (Figure 8.15), but also a mobile client in Android was developed to access the SPs and IdPs. The client, compliant to the SAML ECP profile allows the user to better control the exchange of his identity information.

More details on the code, structures used, modification points, interfaces, configuration, threat analysis and so on can be found in the project deliverables [España Virtual, 2010a] [España Virtual, 2010b].

On the other hand, apart from the validation in the context of the above mentioned project, another type of validation was made through the design, implementation and publication of derived use-cases and architectures based on the ideas presented here. Initially, the architecture and main related concepts were outlined in [Cabarcos, 2009], [Arias et al., 2009] and [Almenárez et al., 2009]. Furthermore, the new applications and proposals that reuse the trust modules presented in this thesis are the following:

- In the context of **smart and mobile television**, we proposed FamTV[10] [Arias et al., 2011a] and FedTV [Almenarez et al., 2011]. FamTV is an architecture centered on providing presence-aware personalized TV. But it includes a security layer where

[10] This work received the *Chester W. Sall Award* for the 2nd place best paper in the IEEE Transactions 2011 (http://www.icce.org/index.php/awards2013)

the trust-based federation will be used to dynamically cope with the huge ecosystem of services and applications that can be accessed from the TV. In turn, FedTV presents an architecture for an enhanced SAML-based mobile client that supports the establishment of dynamic federations to allow cooperation in mobile DTV scenarios (content sharing, service delegation, etc.)

- In the context of **Cloud Computing**, we combined the dynamic federation architecture layers with privacy improvements [Sanchez et al., 2012] to define a federated identity management solution for cooperation, on-demand resources provisioning and delegation in cloud-based scenarios. Furthermore, we designed and prototyped SuSSo [Arias et al., 2012c], an architecture for moving sessions across devices guaranteeing cloud service continuity. The aim with this architecture is to provide users with a mechanism to seamlessly move sessions of services that are given by cloud-based providers (e.g., video or audio streaming). Both proposals build on the dynamic federation architecture presented in this thesis.

During the implementation and the process of new applications development that make use of the architecture modules, we faced different problems and came to interesting conclusions. The main lessons learned through this phase, as well as the remarkable conclusions and limitations found can be summarized in:

- The integration between SP and IdP implementations, even if supporting the same set of IdM specifications, is not an easy task. This fact confirms that the introduction of automation features is important.

- The designed trust logic is easy to integrate with current open source identity federation toolkits. In our tests, for the sake of simplicity, we directly modified the source code of the used toolkits (Authentic and Zxid) in order to include the new functionality. However, the programming of the proposed modules as external APIs to be used from every SP/IdP implementation would be more appropriate and thus is proposed as a future working line derived from this thesis.

- The implemented prototype proves that the core principles of the architecture are workable. Thus, the reputation protocol is realizable and can be included in SAML-based frameworks, and the static operation flow of providers can be modified to provide flexibility in the decisions. In the implementation, tests were focused in the generation, sending, parsing and usage of reputation messages over SAML, thus using a TTL parameter equal to 1. Further validation is performed by means of simulation in section 8.3.

In conclusion , we covered the validation of the architecture proposed in this thesis through the described implementation. As a future step, it would be interesting to test the architecture in a real world FIM environment to validate not only if the architecture is feasible, well-designed and meets the specified requirements, but also if it is useful in real life scenarios (speed, correct decision-making, etc.).

8.5. Conclusions

In this chapter we have presented the work carried out to validate the ideas presented in this dissertation. Validation has been performed through the implementation and testing of prototypes, simulation experiments, behavioral testing of theoretical models, as well as through the dissemination and publication in journals and conference, as well as through contribution to R&D projects.

After all these steps we conclude that the ideas in this thesis are feasible, but also that:

- The architecture can be further tested through its deployment in real world scenarios.
- Validating risk assessment approaches, as pointed out in [Fenz and Ekelhart, 2011], is a complex task. In general, successfully implementing a risk assessment mechanism depends on trust in the gained results. Relying on an incorrect model will result in incorrect data and security decisions. Here we have proved the correctness of our model, but also propose extended validation as future work.
- In the validation of the reputation-based trust model it is required an ecosystem with a high number of services, which forced us to rely on simulations. In this regard, there is a lack of public information on how FIM networks are interconnected, which would have been useful in modeling.

Chapter 9

Conclusions and future lines

The best way to predict the future is to invent it.

Alan Kay, 1971

Contents

9.1. Main Contributions

9.1.1. Technical Contributions

The main technical contributions of this thesis towards the fulfilling of the goals presented in the introductory chapter in order to achieve dynamic federation are detailed below:

1. **Study of FIM specifications and gap analysis.**

   Federated Identity Management (or FIM) recently emerged as a new concept for handling online identity. The main idea behind it is to enable the portability of identity information across otherwise autonomous security domains, allowing the linkage of a person's electronic identity data stored across multiple distinct identity management systems. Based on these premises, FIM enables use-cases such as the widely known SSO, that allows users to authenticate at one site and gain access to multiple services in other domains; but also more rich and complex use-cases, such as attribute sharing, cross-domain user account provisioning, linking of accounts of the same user at different providers, etc.

Due to the importance of the FIM paradigm, the industry and research community have produced a number of standards and specifications representing the fundamental building blocks to accomplish identity federation. However, FIM is still a very new technology (only around a decade old) and so a number of gaps and challenges remain open and need to be filled and solved in order to achieve maturity and wide-scale deployment. In order to identify these gaps, chapter 2 reviewed the state-of-the-art on FIM technologies. This summary is the basis for the gap analysis in Chapter 3, where the main research challenges are described. More specifically, there are important issues with regard to trust, security, privacy, interoperability, assurance, liability and discovery (see details in Chapter 3).

In this dissertation, we center on the trust problem. Trust is a fundamental issue to address scalability. Moreover, the flexibility of every federation framework is tied to the underlying trust model, often poorly defined or even out of the specifications scope. Currently, there are two common approaches in FIM: the *"accept-all-comers"* model, which means that entities always trust every other unknown entity to share identity data; and the rigid trust model, which means that no other entity will be considered trusted, apart from those with pre-configured relationships. In the first case the are obvious security implications, whereas the second case provides security by sacrificing flexibility.

Since there is not a way to establish a trust relationship in a dynamic and secure fashion, we decided to tackle this problem and define new enhanced techniques to achieve ad-hoc dynamic federation. The proposal is based on the introduction of a risk management model and the flexibilization of trust management by including reputation data. The combination of both aspects allows the moving from binary decision-making based on digital certificate lists, towards a more granular model.

2. **Study of trust aspects in FIM.**

   In order to acquire the necessary background to design the dynamic trust model, we started by studying how trust aspects are taken into account in FIM nowadays. Hence, Chapter 2, provides an overview of trust and reputation models; and summarizes related work being carried out by individual researchers, international research projects and organizations involved in standardization. Then, after this objective review, Chapter 3 summarizes the main points of the related work and how it compares to our proposal.

   According to the literature, the inclusion reputation has proven to be useful to enhance security and facilitate interaction among strangers in distributed networks. But in the particular context of FIM, a few proposals exist, being all of them reliant on maintaining centralized information, but not completely distributed.

3. **Study of risk aspects in FIM.**

   In the state-of-the-art revision in Chapter 2, we have also summarized the related work in regard to risk. As documented in the mentioned chapter, a number of formal methodologies have been developed in the IT field. However, in the specific FIM context, there are no defined methods for risk assessment. The only frameworks that provide guidance in this sense are those based on level of authentication assurance (or LoAs). The LoA refers to the degree of confidence in identifying an entity to whom a credential was issued, and the degree of confidence that the entity using the credential is indeed the entity that the credential was issued to. And this concept is used in this way: the higher the risk, the higher the required LoA that must be implemented to allow the transaction. Nevertheless, no complete model exists that considers all the aspects (apart from LoAs) that influence the risk in a FIM transaction. Also, mathematical means to aggregate risk components and provide a final representative value that could be used in decision-making is missing and highly desirable for automation.

   Furthermore, works that integrate both trust and risk data into a model for trust relationship establishment are scarce in IT, and nonexistent in FIM.

   Consequently, the main appeal and contribution of this thesis is the design of an architecture that includes and combines trust and risk to achieve dynamic federation establishment, as well as the definition of the mathematical models for reputation-based trust calculation, risk calculation and its aggregation.

4. **Design of an extended architecture for dynamic federation.**

   Once the gaps were identified and the problem of dynamic federation properly defined, Chapter 4 contributes with the design of the required architecture to fulfill our goals. It is clear that current FIM architectures are limited to provided secure and dynamic means to establish relationships between providers. Thus, based on the general architectural model that is common to FIM systems, we introduce our extensions to extend its functionality. The architecture is composed of a set of logical modules that separate and encapsulate the functionalities required to achieve dynamic federation. The pillars of the architecture are the risk and trust components, which constitute the main contribution of the thesis. The mathematical models implemented by each part are later developed in the subsequent chapters.

   In conclusion, the extension of the architecture satisfies the intended goals, since it makes possible to minimize pre-configuration requirements by allowing relationships to be established on demand based on trust and risk analysis. In order to validate the feasibility of the architecture, we have developed a proof-of-concept application, which is presented in Chapter 8.

5. **Formalization of a computational trust model based on reputation.**

   Having a formal model to compute and represent trust as a number provides a basis for easy implementation and automation. With this premise as a foundation, we developed the trust model that is implemented by the *Trust Manager* component of the architecture for dynamic federation proposed in Chapter 6. It captures the features of current FIM systems and introduces new dimensions to add flexibility and richness. The model includes the definition of a trustworthiness metric, detailing the evidences used, and how they are combined to obtain a quantitative value. Basically, authentication information is merged with behavior data, i.e., reputation or history of interactions. With this, we move from the currently used binary-based decision model into a model that allows more granularity. Trust is computed as a continuous numeric value and continually adjusted taking into account the behavior of the entities. An important part of the model is the inclusion of reputation data, which is a new research line in FIM that opens the door for further investigation. In order to include reputation data in the model we contributed with:

- the definition of a generic protocol to exchange reputation information between FIM entities,
- the definition of the SAML syntax for the reputation protocol messages, and
- the implementation of a prototype and a set of simulation experiments (see Chapter 8) to test the feasibility and benefits of including reputation

Furthermore, our proposal aims to combine trust with risk assessment in order to get a better informed decision-making procedure, which leads to the following contribution and the most important and novel of the thesis.

6. **Formalization of a risk assessment model.**

The proposed risk model that is to be included by participants in federated identity management scenarios is described in Chapter 5. The general goal of the model is to provide a meaningful numerical value that condensates risk information, allows entities to include subjective preferences according to their interests, and aids entities in decision-making. Since a pure quantitative analysis is not possible due to the lack of available statistical data to build the model, we follow a semi-quantitative approach. Qualitative scales are thus required, but instead of staying purely qualitative, the mapping of these scales to quantitative values permits the automation needed for dynamic federation.

The methodology employed to define the risk model consists of three steps. Firstly, we design a taxonomy to capture the different aspects of a relationship in identity federation that may contribute to risk. This approach allows us to decompose the complex problem of risk assessment and to acquire a detailed knowledge. Secondly, based on the taxonomy and aiming at developing a computational model, we propose a set of metrics as a basis to quantify risk. More specifically, metrics are identified in regard to security (cryptographic algorithms and protocols in place, authentication assurance mechanisms, etc), interoperability (of legal, technical and operational policies), knowledge (related to the information known about a collaborator), history of interactions and service specific metrics. Finally, we describe how to aggregate the metrics into a meaningful risk figure, coming to the final formal definition of the model. To develop this part it was required to study multicriteria decision problems and applicable aggregation methods. After the analysis, we decide to use the Multiattribute Utility Theory (MAUT) model, which has been applied and adapted to define the risk aggregation procedure.

7. **Mechanism for decision making based on trust and risk values.**

   Once the models to obtain trust and risk values are defined, the next contribution (developed in Chapter 7) is the definition of a mechanism to aggregate both values and output a decision whether to continue with a transaction or not.

   In the traditional FIM model entities were added and removed as entries in trust lists only as a result of explicit administrative action, reflecting changes to agreements with direct partners. The aim here is that these operations are performed automatically.

   The decision mechanism combines the trust and risk values associated to a transaction, and generates a final value that represents the decision. It is based on a fuzzy aggregation system, that is parametrizable, flexible, allows to model complexities in the trust-risk relationship and captures the subjectivity of entities.

8. **Study of multicriteria decision problems**

   During the definition of the computational trust and risk models we faced the problem of how to combine criteria of different nature, measured on different scales into a meaningful value that can be used in decision making. This problem is known as multicriteria decision making or MCDM.

   There are two main approaches of MCDM, namely 1) the MAUT approach; and 2) the preference modeling approach, whose maximum representative method is the Analytical Hierarchy Process (AHP) [Saaty, 1990]. In MAUT, an absolute score is given to each alternative with respect to each criterion, and the global score, taking into account all the criteria, is obtained by aggregating all the partial scores. By contrast, in preference modeling, a preference degree is assigned to every ordered pair of alternatives, with respect to each criterion. Then, a global preference degree is obtained by aggregating all the partial preference degrees. Similarly, another interesting approach comparable to preference modeling is the usage of MDS techniques [Borg and Groenen, 2005], which allow to rank entities based on the calculation of their similarities or dissimilarities regarding different dimensions. In fact, MDS-based approaches have proven to be useful in security-related problems, such as the selection of the most appropriate peer to interact with based on a set of criteria (cost, distance, security) [Díaz Sánchez, 2008] [Díaz et al., 2006].

Like MAUT, AHP is a compensatory optimization approach. However, AHP uses a quantitative comparison method that is based on pair-wise comparisons of decision criteria, rather than utility and weighting functions. All individual criteria must be paired against all others and the results compiled in matrix form, implying complex calculations of Eigen values and Eigen vectors [Aldrich, 2006]. Apart from complexity, the AHP approach always requires a minimum of two alternatives, and the semantics of the values are relative degrees of preference. In our case we needed a method that allows to give absolute scores, since different two type of decisions are to be made: deciding among several providers, but also deciding whether a provider is suitable to transact or not. In this later case, there are no alternatives, so AHP would not be applicable. The same issue arises with the usage of MDS techniques. For these reasons, we have chosen MAUT as the approach to combine the different dimensions of risk, as well as to combine trustworthiness and risk into the final value for decision-making.

9. **Study of Fuzzy-based aggregation in MCDM.**

Another aspect of modeling MCDM problems is the aggregation of criteria. There are several mathematical functions that can be applied to obtain a final value and the selection of these functions must be carefully performed depending on the features of the problem. The weighted linear aggregation is the simplest and more commonly used method, but its applicability is limited to cases where criteria are independent. However, for interacting criteria, more complex functions are required. Based on these arguments, we have used linear aggregation to combine the risk dimensions, since they are independent. But for the case of combining trust and risk, we have used fuzzy aggregation functions, since linear functions are not able to model the interaction between these two criteria.

Thus, for designing the fuzzy aggregation mechanism, we firstly studied the two main existing alternatives: Mamdani-type fuzzy systems and Sugeno-type fuzzy systems. The most fundamental difference between Mamdani-type FIS and Sugeno-type FIS is the way the crisp output is generated from the fuzzy inputs. While Mamdani-type FIS uses the technique of defuzzification of a fuzzy output, Sugeno-type FIS uses weighted average to compute the crisp output. This means that Mamdani FIS has output membership functions whereas Sugeno FIS has no output membership functions. These features make Mandani a more interpretable and intuitive option, reason for what this type is the most commonly implemented. On the negative part, the computational performance is better in Sugeno. We decided to use Mandani because of its simplicity and the possibility of defining an output with different membership functions that can be mapped to different degrees of cooperation. And the computational cost remains low in this case because the number of rules is small.

10. **Validation**

    Finally, we validated the main contributions in Chapter 8.

    Section 8.2 in the validation chapter (Chapter 8), refines the risk model and completes its definition through a series of formal analytical tests. Validation is conducted with the aim to demonstrate that the risk model works in conformance to the associated principle guidelines, that the output is correct. More specifically, we test how the model is capable of handling the set of quantitative metrics defined and use them as an input to generate the associated risk value. We show that the final risk value:

    - is relative to the perception, assets and needs of the provider that is making the evaluation,
    - provides information on the assurance level coverage, and
    - can be used in decision making. The final value allows to discard those entities that do not satisfy minimum requirements; and also to make a comparative ranking of entities when there are several options available

      Based on the performed tests, recommendations are given on how to implement aggregation. Finally, a detailed catalogue summarizing the set of defined metrics is contained in Appendix B.

Furthermore, the reputation-based trust model was validated through a simulation-based testing of the protocol. We showed the increasing cooperation rates when incorporating reputation data with respect to the pre-configured model, and how the reputation information allows to identify and isolate bad entities. Further work could be done specifically in regard to this part of the thesis, as it will be pointed out in section 9.3.

In turn, the architecture was validated through the implementation of a prototype and its deployment in the context of a national R&D project. Also, different modules of the architecture were reused as part of other works published in [Arias et al., 2012c] [Arias et al., 2011a] [Almenarez et al., 2011], that illustrate use-cases based on dynamic federation.

9.1.2. Other Contributions

Apart form the technical contributions described above, another kind of activities were also carried out in the context of this dissertation that convey an added value to its realization. More specifically:

1. **Dissemination.**

   The dissemination tasks consisted on the publication of papers and contribution to conferences and journals where the main ideas of the thesis were subject to peer review, evaluation and discussion. The main publications are detailed in Chapter 1, section 1.3.

   Furthermore, part of the contents in this thesis were developed as a research line in two national R&D projects: "España Virtual"[1] and CONSEQUENCE[2]. Both projects included an specific working package for "*Security and Identity Management*", where our ideas on dynamic federation were contributed. In addition, during a research stay at NEC Laboratories Europe, the work on the integration of reputation in FIM and the simulations presented in Chapter 8 were included in the deliverables of an internal business project centered in IdM.

2. **Identification of new research lines.**

[1] http://www.espanavirtual.org/

[2] http://consequence.it.uc3m.es

In section 9.3, considering the limitations and the points of improvement of our ideas, we identify a set of open issues for future research works.

9.2. Conclusions

After all the work carried out during the development of the thesis, we can conclude that:

- FIM technologies are a solution to the problem of identity portability and sharing of identity data across different domains. However, the underlying trust models existing nowadays limit their applicability.
- There are currently two main options to interact in a FIM context: to only interact with know entities that are pre-configured to be trusted (by means of complex contractual frameworks); or to interact with every entity independently of it is known or unknown. In the first case, the initial setup complexity is a high barrier and may not worth adopting these procedures for a short-term collaboration because time and cost will probably not outbalance the rewards of cooperation. In the second case, security problems may arise.
- The architecture proposed in this thesis solves the above problems and enables the establishment of dynamic secure federations by including more flexible trust management and risk assessment. The new trust mechanisms allow to dynamically acquire data (i.e., reputation) about previously unknown parties and compute a trust value. The new risk mechanisms allow to analyze the features of other entities and determine a quantitative number representing the risk involved in transacting. Both values are combined to output a decision. They are also updated and computed on a per-transaction basis, adapting the decisions to the real behavior of the entities and to the value of the transaction in course.
- The proposed trust model provides a quantitative formalization that captures current FIM trust features and introduces flexibility. The main features of the model are:
    - Trust is based on two kind of evidences: authentication trust and behavior trust. Authentication trust is based on digital certificates, and behavior trust is based on the history of interactions and/or on reputation data. The inclusion of reputation data implies also the definition of a protocol for the exchange of

  reputation in FIM scenarios. The difference with the closest related works that include reputation is that they use centralized storage of data, while our model is completely distributed. However, our approach is not competing but integrative, i.e., different algorithms for the calculation and update of the behavior trust can be used.

  - The protocol for reputation acquisition is outlined as an initial simple model based in the existing reputation protocols designed for P2P systems, but defining the specific features to be applicable in FIM, i.e., the data that should be included and also how it can be implemented over a concrete specification (SAML).

- The inclusion of risk assessment into the decision-making process is the main novelty in our proposal, which constitutes the first contribution of this kind in the context of FIM. Though the concept of including risk is suggested in the literature, there are no works that explicitly define a model to compute it.

  The risk assessment model derives from a taxonomy that we have designed to understand the federation procedure and to categorize the different aspects of risk. Based on the taxonomy, we defined a set of metrics including the semantic description, the procedure to obtain them, and the associated qualitative scales. Next, we provided a quantitative mapping of the metrics and we put forward a collection of mathematical formulas that permit the aggregation of the metrics into a final risk value. This value has proven useful to provide rapid information about whether the potential cooperating entity satisfies the risk policies of the evaluating entity and to which degree, constituting a perfect element for decision-making. However, as we will see in section 9.3, the validation of the model can be extended.

- The implementation and deployment of a full prototype that follows the architecture for dynamic federation is a challenge. Firstly, there is a need for common formats and standards to convey all the information used in the trust and risk computation (see section 9.3). But also, a real deployment raises many issues. For example, for the reputation system to be realizable it is required that the protocol is implemented by a high number of providers (or by providers with high connectivity), otherwise the benefits would be low, since it would be hard to find reputation data about potential cooperators. Further validation is required, as we identify in the next section, in order for the solution to be fully realizable.

- The achievement of dynamic federation will enable the adoption of FIM at a wide scale and make possible a new range of use-cases. A good example is the applicability to the paradigm of Instant Virtual Enterprises or IVEs, where a set of partners must quickly configure themselves to exploit a a concrete business opportunity. In this context, a rapid mean to federate the identity management systems of the co-operating partners is required, and current FIM technology does not satisfy these high demands of agility.

Summarizing, we have developed an initial approach towards the realization of dynamic federation that is based on improving the process of trust relationship establishment. As the Internet Society (ISOC) remarks, *"the issue of trust is both important and crucial for the long-term growth and success of the Internet. There is no debate about the explosive innovation that has occurred as a result of building the Internet and, if the promise of federated identity can be realized, a similar explosion in innovation will occur."* We believe indeed that dynamic federation will enable a wide new range of possible business and models of cooperation. However, our proposal is a set of preliminary ideas, partially validated and, some of them, prototyped. As a preliminary work, many limitations and weaknesses exist that open the room for further research and improvement. With the aim to identify these lacks and open issues, the following section presents the main future research lines that can be derived from here.

9.3. Future Research Lines

In this thesis we have contributed to evolve federated identity management towards more flexible, dynamic and secure models. However, there are still open issues and future lines that require further research to make dynamic federation a reality. We recognize several areas where the architecture and components presented here can be improved or extended, all of them explained below:

1. **Extension of FIM metadata documents to convey more information.**

   The risk model proposed in this thesis is based on a set of defined metrics belonging to the following categories: security, interoperability, knowledge, service specific and historical interactions. The procedures to extract the metrics values consist on analyzing the information available for their category, assign them a qualitative scale and then map it to a quantitative value. For example, for the security metric "confidentiality at message level" it is required to obtain the algorithms/protocols used for information encryption in order to determine the associated assurance level according to specific features, such as strength, key size, etc. Hence, the information required for extracting the metrics must be expressed and exchanged between entities. However, there are no standard formats and conventions to convey all the required data (we developed our risk model assuming that this information was available).

Currently, there is support in SAML for communicating the supported cryptographic algorithms through metadata extensions that are being defined in an OASIS draft [Cantor, 2010]. And there is also support for communicating levels of authentication assurance in SAML. But there are no means to transmit other kind of security data. For the rest of categories, it is also required to express information such as the laws under which the provider operates, details of operational policies, service parameters, etc.

Thus, an interesting future line would be to define the extensions to the metadata documents that allow to express the information required to derive all the risk metrics. In this regard, and following our integrative philosophy, the work that is being carried out by the OIX organization (see Chapter 2) could be linked and used in the extension of metadata information.

2. **Definition of detailed frameworks for the risk metrics.**

   Apart from the extension of the metadata documents to include all the information required for calculating the risk metrics, another closely related line for further research could be the definition of more detailed frameworks (or new ones) for some of the risk metrics.

   In this thesis we defined a set of risk metrics together with the procedure to obtain them. For a number of metrics (e.g., for those related with cryptography) there is extensive expert knowledge available in the literature. In those cases, we used this knowledge as a basis for defining detailed frameworks to assign the qualitative scales. However, in other cases, this knowledge base does not exist, so we defined simpler frameworks that are just a starting point and do not cover all the nuances and complexities. Hence, it is required to complete the process of investigating all the aspects that may contribute to the assignation of a metric to a specific category and generate this expert knowledge to complete the definition of better frameworks.

3. **Definition of a SLA standard format and SLA negotiation for federated identity management.**

   A SLA is put in place for the establishment of federations whether explicitly or implicitly. Here we assume that entities have public SLAs where information about the service provision features can be found and used for risk and trust calculation. In this regard, a future research line could be the definition of standards formats for SLAs in the context of FIM. It is required to define all the parameters involved in services provided by IdPs/SPs for an easier automation and usage in the establishment of dynamic federations.

   Furthermore, we contemplate the possibility of setting different SLAs depending on the initial risk/trust that is placed on an entity and change them (giving or denying more privileges) according to how the relationship evolves over time. For this purpose, a SLA negotiation approach can be introduced and investigated as a future line. In this regard, the European research project SLA@SOI [SLAatSOI Project, 2011] focused on the process of negotiating SLAs and provisioning, delivery and monitoring of services for highly dynamic and scalable service consumption. The methods researched in this project could thus be integrated with the work presented here.

4. **Further study and analysis of the reputation protocol.**

   Developing an implementable reputation system is an art involving many separate design problems and choices. We have defined the main features for the reputation protocol to be applicable in FIM: the data that should be included and also how it can be implemented over a concrete specification (SAML). We started a simulation testbed to prove the benefits of the protocol, showing that the cooperation rates improve and that the bad entities are isolated. But further work is still required. Among others, the following aspects could be addressed:

   - Study and comparison of different forwarding strategies. The initially proposed protocol consists on forwarding reputation requests to every trusted node in the DTL. More intelligent strategies could be designed (e.g., identifying those nodes with more connections and forward only to them) to improve the performance of the protocol.

- Security analysis and enhancement. According to the initially proposed protocol, if the reputation queries are forwarded beyond direct neighbors, the querying entity may receive reputation response messages form unknown entities. It could lead to attacks such as the injection false response messages. This, and other possible attacks should be simulated and investigated in order to improve the protocol to take into account high security considerations.

5. **Additional validation and evaluation.**

Apart from the further study suggested above in regard to the reputation protocol and its complete validation, both the risk model and the general architecture need also to be subject of an extended validation.

The risk model has been validated by testing if it conforms to the design objectives. To validate results in the context of risk assessment, apart from testing that the model is correct according to the design requirements, researchers have proposed other three main approaches: 1)using experts, 2) using alternate decision processes, and 3) using statistical evidence.

In our case, since there are no other risk-based decision tools in FIM, neither threats statistical data are available, the validation using experts would be the next step in validation.

The process would consist on running our model in IdPs/SPs deployed in the real world and comparing the decision output with the decision of the administrator.

6. **Study and definition of business models based on dynamic federation.**

A big issue in the adoption of federation, as for every other technology, is business. Recent studies point out that one of the main reasons that the wide adoption of federation is not happening is because service providers do not have sufficient incentives to become relying parties. There is a need to define models for monetizing identity services, investigate business needs and define models where the benefits for all the involved parties are balanced.

7. **Integration of the dynamic federation in different specifications.**

   The main goal in this thesis was to define an infrastructure generic enough to be applicable to any federation specification. Hence, we studied the main documents of the different FIM protocols as a basis for the definition. However, whenever particularization was required to go deeper on the description of the model, we based on the SAML specifications. Also, the developed prototype was implemented over SAML. Further work is required to integrate and implement the proposal in the rest of the specifications.

Appendices

Appendix A

Glossary

For the purposes of the present document, the following abbreviations apply:

A

ACI Assurance Compliance Index

AHP Analytical Hierarchy Process

API Application Programming Interface

ARP Attribute Release Policy

ADL Architecture Description Language

B

BAL Business Anchor List

B2B Business to Business

C

CA Certification Authority

CENIT Consorcio Estratégico Nacional de Investigación Técnica

CoT Circle of Trust

D

DTL Dynamic Trust List

E

ECP Enhanced Client Proxy

ETSI European Telecommunications Standards Institute

F

FIM Federated Identity Management

FIS Fuzzy Inference System

H

HTTP Hypertext Transfer Protocol

I

IETF Internet Engineering Task Force

IdP Identity Provider

IMS Internet Multimedia Subsystem

ISO International Organization for Standardization

ITU-T ITU Telecommunication Standardization Sector

IVE Instant Virtual Enterprise

J

JCR Journal Citation Report

L

LDAP Lightweight Directory Access Protocol

LoA Level of Authentication assurance

M

MAUT Multi-Attribute Utility Theory

MCDM Multi-Criteria Decision Making

MDS Multi Dimensional Scaling

N

NGN Next Generation Networks

O

OASIS Organization for the Advancement of Structured Information Standards

OIX Open Identity Exchange

OHSAS Occupational Health and Safety Assessment Series

P

PKI Public Key Infrastructure

P2P Peer To Peer

R

REFEDS Research and Education Federations

RFC Request For Comments

RP Relying Party

S

SAML Security Assertion Markup Language

SDO Standards Developing Organization

SLA Service Level Agreement

SLA Single Log Out

SOAP Simple Object Access Protocol,

SP Service Provider

SSL Secure Sockets Layer

SSO Single Sign On

SSTC Security Services Technical Committee

STS Security Token Service

T

TAL Trust Anchor List

TERENA Trans-European Research and Education Networking Association

TTL Time To Live

U

UML Unified Modeling Language

V

VE Virtual Enterprise

VO Virtual Organization

X

XML eXtensible Markup Language

Appendix B

FIM Risk Metrics Catalogue

In this Appendix, we summarize the set of metrics for risk quantification in FIM proposed in this thesis based on the taxonomy in 5.1. To simplify the presentation of the metrics, they are classified and grouped according to different features, namely:

- **General Category**, indicates if the metric is under the *Pre-Federation* or the *Post-Federation* category.
- **Level**, indicates the taxonomic level in which the metric is located.
- **Type**, indicates if the metric is *Basic* or *Aggregated.*
- **Formula**, indicates the mathematical expression for calculating the metric value. Only for *Aggregated* metrics.
- **Assurance Framework**, specifies features about the assurance framework used to define the metric. It can be whether a well-known framework or a self-defined framework. In the first case, if alternative assurance frameworks are available, they are also indicated. In the latter case, it is also specified if the framework is based on expert recommendations or defined in a higher abstract level.
- **Assurance Scale**, indicates the qualitative scale used by the assurance framework, as well as the quantitative mapping used fro aggregation.

The catalogue encompasses 15 *Basic* and 7 *Aggregated* metrics in the *Pre-Federation* category; and 12 *Basic* metrics and 5 *Aggregated* in the *Aggregated* category. All the metrics are summarized below.

B.1. Pre-Federation Metrics

Table B.1 shows the set of basic metrics in Pre-Federation.

Level	Metric	Definition	Scale	Framework
L4	CONF_{TL}	Confidentiality assurance at the transport level	3-level scale {Low,Medium,High} with associated quantitative values {1, 2, 3}	Self-defined based primarily on NIST recommendations and SSL/TLS specifications
	CONF_{ML}	Confidentiality assurance at the message level		
	INT_{TL}	Integrity assurance at transport level		
	INT_{ML}	Integrity assurance at message level		
	AUTH_{TL}	Authentication assurance at transport level		
	AUTH_{ML}	Authentication assurance at message level		
L3	NON-REP	Non-repudiation assurance		Self-defined, high abstract level
	AV	Availability assurance		
	ACC	Accountability assurance		
	LOP	Level of Protection, privacy assurance		
	$\text{INTEROPT}_{\text{T}}$	Technical interoperability assurance		
	$\text{INTEROP}_{\text{O}}$	Operational interoperability assurance		
	$\text{INTEROP}_{\text{L}}$	Legal interoperability assurance		
	KNOW_{D}	Direct knowledge assurance	2-level scale {False, True} with associated quantitative values {0,1}	
	KNOW_{I}	Indirect knowledge assurance		

Table B.1: Basic Pre-Federation Metrics

In turn, Table B.2 shows the set of aggregated metrics in the Pre-federation category. The formulas to compute the aggregated metrics are the same for all the metrics (based on the MAUT technique [Keeney and Raiffa, 1993]). The difference lies on the score vectors (SV) and the reference vectors (RV) employed in each case. Thus, the score vectors for each metric are shown in Table B.3. The RVs, however, will depend on the risk policies defined by the entity applying the risk model.

Level	Metric	Definition	Formula
L3	CCONF	Constrained Confidentiality Assurance	$CAgg = \begin{cases} Agg\ \ if\ \ ACI = 1 \\ 0\ \ if\ \ ACI \neq 1 \end{cases}$, where $ACI = \begin{cases} 1\ \ if\ \ SV \geq RV(i)\ \ \forall\ \ i \\ \dfrac{\lvert \cup SV \rvert}{n}\ \ otherwise \end{cases}$
	CINT	Constrained Integrity Assurance	
	CAUTH	Constrained Authentication Assurance	
L2	CSA	Constrained Security Assurance	
	CIA	Constrained Interoperability Assurance	
	CKNOW	Constrained Knowledge Assurance	
L1	PreFedA	PreFederation Assurance	

Table B.2: Aggregated Pre-Federation Metrics

Metric	**Score Vector**
CCONF	[|| CONF_{TL} ||, || CONF_{ML} ||]
CINT	[|| INT_{TL} ||, || INT_{ML} ||]
CAUTH	[|| AUTH_{TL} ||, || AUTH_{ML} ||]
CSA	[CCONF, CINT, CAUTH, || NON-REP ||, || AV ||,|| ACC ||, ||LOP ||]
CIA	[|| $\mathrm{INTEROP}_{T}$||,|| $\mathrm{INTEROP}_{O}$||,|| $\mathrm{INTEROP}_{L}$||]
CKNOW	[|| KNOW_{D}||,|| KNOW_{I} ||]
PreFedA	[CSA, CIA, CKNOW]

Table B.3: Score Vectors for Aggregated Pre-Federation Metrics

B.2. Post-Federation Metrics

Table B.4 shows the set of basic metrics in Post-Federation.

Level	**Metric**	**Definition**	**Scale**	**Framework**
L4	$\mathrm{CONF}_{\mathrm{TL}}$	Confidentiality assurance at the transport level	3-level scale {Low,Medium,High} with associated quantitative values {1, 2, 3}	Self-defined based primarily on NIST recommendations and SSL/TLS specifications
	$\mathrm{CONF}_{\mathrm{ML}}$	Confidentiality assurance at the message level		
	$\mathrm{INT}_{\mathrm{TL}}$	Integrity assurance at transport level		
	$\mathrm{INT}_{\mathrm{ML}}$	Integrity assurance at message level		
	$\mathrm{AUTH}_{\mathrm{TL}}$	Authentication assurance at transport level		
	$\mathrm{AUTH}_{\mathrm{ML}}$	Authentication assurance at message level		
L3	NON-REP	Non-repudiation assurance		Self-defined, high abstract level
	AV	Availability assurance		
	ACC	Accountability assurance		
	LOP	Level of Protection, privacy assurance		
	LOA	Level of Authentication assurance	4-level scale {Little, Some, High, Very High} with associated quantitative values {1,2,3,4}	NIST 800-63 [Nadalin et al., 2006]. Other alternative availabel frameworks are ATSC2[1], ATSC3[2], STORK[3]
	HINT	Historical interactions assurance	3-level scale {Low,Medium,High} with associated quantitative values {1, 2, 3}	Self-defined

Table B.4: Basic Post-Federation Metrics

In turn, Table B.5 shows the set of aggregated metrics in the Post-Federation category. The formulas to compute the aggregated metrics are the same for all the metrics (based on the MAUT technique [Keeney and Raiffa, 1993]). The difference lies on the score vectors (SV) and the reference vectors (RV) employed in each case. Thus, the score vectors for each metric are shown in Table B.6.

Level	Metric	Definition	Formula
L3	CCONF	Constrained Confidentiality Assurance	$CAgg = \begin{cases} Agg\ if\ \ ACI = 1 \\ 0\ \ if\ \ ACI \neq 1 \end{cases}$, where $ACI = \begin{cases} 1\ \ if\ \ SV \geq RV(i)\ \ \forall\ \ i \\ \frac{\lvert \cup SV \rvert}{n}\ \ otherwise \end{cases}$
	CINT	Constrained Integrity Assurance	
	CAUTH	Constrained Authentication Assurance	
L2	CSA	Constrained Security Assurance	
L1	PostFedA	PostFederation Assurance	

Table B.5: Aggregated Post-Federation Metrics

<table>
<tr><th>Metric</th><th>Score Vector</th></tr>
<tr><td>CCONF</td><td>[|| CONF$_{TL}$ ||, || CONF$_{ML}$ ||]</td></tr>
<tr><td>CINT</td><td>[|| INT$_{TL}$ ||, || INT$_{ML}$ ||]</td></tr>
<tr><td>CAUTH</td><td>[|| AUTH$_{TL}$ ||, || AUTH$_{ML}$ ||]</td></tr>
<tr><td>CSA</td><td>[CCONF, CINT, CAUTH, || NON-REP ||, || AV ||,|| ACC ||, ||LOP ||]</td></tr>
<tr><td>PostFedA</td><td>[CSA, || LOA ||, || HINT ||]</td></tr>
</table>

Table B.6: Score Vectors for Aggregated Post-Federation Metrics

[1] www.ref.gv.at/AG-IZ-Sicherheitsklassen-Sec.1719.0.html

[2] www.ref.gv.at/Sicherheitsklassen.2329.0.html

[3] www.eid-stork.eu

Bibliography

[Adams, C. and Farrell, S., 1999] Adams, C. and Farrell, S. (1999). Internet X.509 Public Key Infrastructure Certificate Management Protocols.

[Aldrich, 2006] Aldrich, J. (2006). Eigenvalue, eigenfunction, eigenvector, and related terms. *Earliest Known Uses of Some of the Words of Mathematics.*

[Almenares et al., 2011] Almenares, F., Marín, A., Díaz-Sánchez, D., and Arias, P. (2011). Personal networks federation in mobile DTV. In *IEEE International Conference on Consumer Electronics (ICCE)*, pages 385 –386.

[Almenarez et al., 2011] Almenarez, F., Arias, P., Díaz-Sanchez, D., Marín, A., and Sánchez, R. (2011). FedTV: personal networks federation for IdM in mobile DTV. *IEEE Transactions on Consumer Electronics*, 57(2):499 –506.

[Almenárez et al., 2009] Almenárez, F., Arias, P., Marín, A., and Díaz, D. (2009). Towards dynamic trust establishment for identity federation. In *Proceedings of the 2009 ACM Euro American Conference on Telematics and Information Systems: New Opportunities to increase Digital Citizenship*, EATIS '09, pages 25:1–25:4.

[Almenárez et al., 2004] Almenárez, F., Marín, A., Campo, C., and García, C. (2004). PTM: A pervasive trust management model for dynamic open environments. In *First Workshop on Pervasive Security, Privacy and Trust PSPT.*

[Ambler and Lines, 2012] Ambler, S.W. and Lines, M. (2012). *Disciplined Agile Delivery: A Practitioner's Guide to Agile Software Delivery in the Enterprise.* IBM Press.

[Anderson, 2004] Anderson, C. (2004). *The long tail.* Wired Magazine.

[Anderson et al., 2000] Anderson, Lorin W., Krathwohl, David R., Airasian, Peter W., Cruikshank, Kathleen A., Mayer, Richard E., Pintrich, Paul R., Raths, James, and Wittrock, Merlin C. (2000). *A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives, Abridged Edition.* Allyn & Bacon, 2 edition.

[Apache, 2012] Apache (2012). The Apache website. `http://www.apache.org/`. [Online; accessed 06-February-2013].

[Arias, 2011] Arias, P. (2011). Risk assessment for better Identity Management in pervasive environments. In *IEEE International Conference on Pervasive Computing and Communications Workshops (PERCOM Workshops)*, pages 389 –390.

[Arias et al., 2012a] Arias, P., Almenares, F., Sánchez Guerrero, R., Marín, A., and Díaz-Sánchez, D. (2012a). Multi-device Single Sign-on for cloud service continuity. In *IEEE International Conference on Consumer Electronics (ICCE)*, pages 644 –645.

[Arias et al., 2009] Arias, P., Almenárez-Mendoza, F., Marín-López, A., and Díaz-Sánchez, D. (2009). Enabling SAML for Dynamic Identity Federation Management. In Wozniak, J.and Konorski, J., Katulski, R., and Pach, A., editors, *Wireless and Mobile Networking*, volume 308 of *IFIP Advances in Information and Communication Technology*, pages 173–184. Springer Boston.

[Arias et al., 2012b] Arias, P., Almenárez-Mendoza, F., Marín-López, A., Díaz-Sánchez, D., and Sánchez-Guerrero, R. (2012b). A metric-based approach to assess risk for "On Cloud" Federated Identity Management. *Springer's Journal of Network and Systems Management, Special Issue on Cloud Computing, Networking, and Service (CCNS) Management.*, 20(4):513 –533.

[Arias et al., 2011a] Arias, P., Guerrero, R.S., Mendoza, F.A., Díaz-Sánchez, D., and Marín López, A. (2011a). FamTV: An architecture for presence-aware personalized television. *IEEE Transactions on Consumer Electronics*, 57(1):6 –13.

[Arias et al., 2012c] Arias, P., Mendoza, F.A., Guerrero, R.S., Marín López, A., and Díaz-Sánchez, D. (2012c). SuSSo: Seamless and Ubiquitous Single Sign-on for Cloud Service Continuity across devices. *IEEE Transactions on Consumer Electronics*, 58(3):1425 – 1433.

[Arias et al., 2011b] Arias, P., Sánchez, R., Almenares, F., and Díaz-Sánchez, D. (2011b). Presence-aware personalized television. In *IEEE International Conference on Consumer Electronics (ICCE)*, pages 769 –770.

[Atwood et al., 2007] Atwood, M., Conlan, R.M., Cook, B., Culver, L., Elliott-McCrea, K., Halff, L., Hammer-Lahav, E., Laurie, B., Messina, C., and Panzer, J. (2007). OAuth Core 1.0. *OAuth Core Workgroup*.

[Authentic, 2007] Authentic (2007). Liberty-compliant Identity Provider. `http://authentic.labs.libre-entreprise.org/`. [Online; accessed 06-February-2013].

[Baldwin et al., 2007] Baldwin, A., Mont, M.C., Beres, Y., and Shiu, S. (2007). On identity assurance in the presence of federated identity management systems. In *ACM Workshop on Digital Identity Management(DIM)*, volume 7.

[Barker et al., 2011] Barker, E., Barker, W., Burr, W., Polk, W., and Smid, M. (2011). *Recommendation for Key Management–Part 1: General (Revision 3)*, volume 800.

[Bass et al., 2003] Bass, L., Clements, P., and Kazman, R. (2003). *Software architecture in practice*. Addison-Wesley Professional.

[Beatty et al., 2004] Beatty, J., Hodges, J., Ellison, G., Kemp, J., Rouault, J., Kainula, J., Wason, T., and Thompson, P. (2004). Liberty ID-WSF Web Services Framework Overview.

[Beliakov and Warren, 2001] Beliakov, G. and Warren, J. (2001). Appropriate choice of aggregation operators in fuzzy decision support systems. *IEEE Transactions on Fuzzy Systems*, 9(6):773–784.

[Bertino et al., 2009] Bertino, E., Martino, L., Paci, F., and Squicciarini, A. (2009). *Security for Web Services and Service-Oriented Architectures*. Springer.

[Bertino et al., 2007] Bertino, E., Squicciarini, A.C., and Bhargavspantzel, A. (2007). Trust negotiation in identity management. *IEEE Security & Privacy*, pages 55–63.

[Bertocci et al., 2007] Bertocci, V., Serack, G., and Baker, C. (2007). *Understanding windows cardspace: an introduction to the concepts and challenges of digital identities*. Addison-Wesley Professional, first edition.

[Black, 2008] Black, P.E. (2008). SAMATE's contribution to information assurance. *NIST Special Publication*, 500(264):2.

[Boeyen et al., 2004] Boeyen, S., Ellison, E.G., Sengodan, S., Shinkar, E.S., and Thompson, C.P. (2004). Trust Models Guidelines.

[Borg and Groenen, 2005] Borg, Ingwer and Groenen, Patrick JF (2005). *Modern multidimensional scaling: Theory and applications*. Springer.

[Boursas and Danciu, 2008] Boursas, L. and Danciu, V.A. (2008). Dynamic inter-organizational cooperation setup in Circle-of-Trust environments. In *Network Operations and Management Symposium, 2008. NOMS 2008. IEEE*, pages 113 –120.

[Boursas and Hommel, 2006] Boursas, L. and Hommel, W. (2006). Policy-based service provisioning and dynamic trust management in Identity Federations. In *IEEE International Conference on Communications, 2006. ICC'06.*, volume 5, pages 2370–2375.

[Boursas and Hommel, 2007] Boursas, L. and Hommel, W. (2007). Derivation and use of trust and risk management parameters in dynamic federated environments. In *Proceedings of the 14th Annual Workshop of HP Software University Association, Leibniz Supercomputing Center, Munich, Germany.*

[Box et al., 2000] Box, D., Ehnebuske, D., Kakivaya, G., Layman, A., Mendelsohn, N., Nielsen, He. Frystyk, Thatte, S., and Winer, D. (2000). Simple Object Access Protocol (SOAP) 1.1. Technical report.

[Bray et al., 2008] Bray, T., Paoli, J., Sperberg-McQueen, C. M., Maler, E., and Yergeau, F. (2008). Extensible Markup Language (XML) 1.0 (Fifth Edition). World Wide Web Consortium, Recommendation REC-xml-20081126.

[Cabarcos, 2009] Cabarcos, P.A. (2009). Dynamic Trust Relationship Establishment in Federated Identity Management (Master Thesis).

[Camarinha-Matos et al., 2005] Camarinha-Matos, L., Silveri, I., Afsarmanesh, H., and Oliveira, A. (2005). Towards a framework for creation of dynamic virtual organizations. *Collaborative networks and their breeding environments*, pages 69–80.

[Camp, 2010] Camp, J. (2010). Identity Management's Misaligned Incentives. *IEEE Security & Privacy*, 8(6):90–94.

[Cantor and (eds.), 2003] Cantor, S. and (eds.), J. Kemp (2003). Liberty ID-FF Protocols and Schema Specification. Version 1.2. Liberty Alliance Project.

[Cantor et al., 2005a] Cantor, S., Hirsch, F., Kemp, J., Philpott, R., and (eds.), E. Maler (2005a). Bindings for the OASIS Security Assertion Markup Language (SAML) v2.0. OASIS Standard.

[Cantor et al., 2005b] Cantor, S., Kemp, J., Philpott, R., and (eds.), E. Maler (2005b). Assertions and Protocols for the OASIS Security Assertion Markup Language (SAML) v2.0. OASIS Standard.

[Cantor et al., 2005c] Cantor, S., Moreh, J., Philpott, R., and (eds.), E. Maler (2005c). Metadata for the OASIS Security Assertion Markup Language (SAML) v2.0. OASIS Standard.

[Cantor, 2010] Cantor, S. (ed.) (2010). SAMLv2.0 Metadata Profile for algorithm support Version 1.0.

[Carnegie Mellon Software Engineering Institute (SEI), 2006] Carnegie Mellon Software Engineering Institute (SEI) (2006). How do you define Software Architecture? `http://www.sei.cmu.edu/architecture/start/glossary/community.cfm`. [Online; accessed 06-February-2013].

[Casare and Sichman, 2005] Casare, S. and Sichman, J. (2005). Towards a functional ontology of reputation. In *Proceedings of the 4th ACM International joint Conference on Autonomous Agents and Multiagent Systems*, pages 505–511.

[Chadwick, 2009] Chadwick, D. (2009). Federated identity management. *Foundations of Security Analysis and Design*, pages 96–120.

[Chernick et al., 2005] Chernick, C. M., Edington III, C., M.J., Fanto, and Rosenthal, R. (eds.) (2005). *Guidelines for the Selection and Use of Transport Layer Security (TLS) Implementations - NIST Special Publication 800-52.*

[Clements et al., 2003] Clements, P., Garlan, D., Little, R., Nord, R., and Stafford, J. (2003). Documenting software architectures: views and beyond. In *25th IEEE International Conference on Software Engineering, 2003. Proceedings*, pages 740–741.

[Clowes and Brathwaite, 2009] Clowes, N. and Brathwaite, L. (2009). D4. 2 Final report on eID process flows.

[Cornelli et al., 2002] Cornelli, F., Damiani, E., di Vimercati, S.D.C., Paraboschi, S., and Samarati, P. (2002). Choosing reputable servents in a P2P network. In *Proceedings of the 11th ACM International Conference on World Wide Web*, pages 376–386.

[Cugini et al., 1997] Cugini, J., Damianos, L., Hirschman, L., Kozierok, R., Kurtz, J., Laskowski, S., and Scholtz, J. (1997). Methodology for evaluation of collaboration systems. *The evaluation working group of the DARPA intelligent collaboration and visualization program, Rev*, 3.

[Cutler, 2007] Cutler, R. (ed.) (2007). Liberty Identity Assurance Framework. Liberty Alliance Project.

[Dabrowski and Pacyna, 2008] Dabrowski, M. and Pacyna, P. (2008). Modular reference framework architecture for identity management. In *11th IEEE Singapore International Conference on Communication Systems (ICCS)*, pages 743–749.

[Damiani et al., 2002] Damiani, E., di Vimercati, D.C., Paraboschi, S., Samarati, P., and Violante, F. (2002). A reputation-based approach for choosing reliable resources in peer-to-peer networks. In *Proceedings of the 9th ACM conference on Computer and Communications Security*, pages 207–216.

[Dang, 2008] Dang, Q (2008). *Recommendation for applications using approved hash algorithms - NIST Special Publication 800-107*. US Department of Commerce, National Institute of Standards and Technology.

[Díaz et al., 2006] Díaz, Daniel, Marín, Andrés, Almenárez, Florina, García-Rubio, Carlos, and Campo, Celeste (2006). Context awareness in network selection for dynamic environments. In *Personal Wireless Communications*, pages 216–227. Springer.

[Díaz Sánchez, 2008] Díaz Sánchez, Daniel (2008). Contribuciones a protocolos y mecanismos de análisis y decisión para control de acceso en entornos distribuidos. tesis doctoral.

[Dierks and Rescorla, 2006] Dierks, T. and Rescorla, E. (2006). RFC 2446: The Transport Layer Security (TLS) Protocol Version 1.1.

[Dierks and Rescorla, 2008] Dierks, T. and Rescorla, E. (2008). RFC 5246: The Transport Layer Security (TLS) Protocol Version 1.2.

[Douceur, 2002] Douceur, J. (2002). The sybil attack. *Peer-to-peer Systems*, pages 251–260.

[E. Tiffany, 2008] E. Tiffany, P. Madsen, S. Cantor (eds.) (2008). Level of Assurance Authentication Context Profiles for SAML 2.0. OASIS Working Draft.

[Eastlake et al., 2012] Eastlake, D., Reagle, Solo, J., D., Hirsch, F., Nyström, M., Roessler, T., and Yiu, K. (2012). *XML-Signature Syntax and Processing Version 1.1.* W3C Working Draft.

[Eastlake et al., 2002a] Eastlake, D., Reagle, J., and Solo, D. (2002a). *XML-Signature Syntax and Processing.* W3C Recommendation.

[Eastlake et al., 2008] Eastlake, D., Reagle, J., Solo, D., Hirsch, F., and Roessler, T. (2008). *XML Signature Syntax and Processing (Second Edition).* W3C Recommendation.

[Eastlake et al., 2002b] Eastlake, D. E., Reagle, J. M., Imamura, T., Dillaway, B., and Simon, (eds.) (2002b). Xml Encryption Syntax and Processing. World Wide Web Consortium, Recommendation REC-xmlenc-core-20021210.

[ENISA, 2006] ENISA (2006). Risk management: Implementation principles and inventories for risk management/risk assessment methods and tools.

[España Virtual, 2010a] España Virtual (2010a). E.6.1.5 Informe de avance en la investigación en Seguridad e Identidad. Entregable del Proyecto CENIT España Virtual.

[España Virtual, 2010b] España Virtual (2010b). Informe de avance en el activo experimental “Georreferenciación semántica del conocimiento”. Entregable del Proyecto CENIT España Virtual.

[ETSI, 2011] ETSI (2011). Identity and access management for Networks and Services; Dynamic federation negotiation and trust management in IdM systems.

[Fenz and Ekelhart, 2011] Fenz, S. and Ekelhart, A. (2011). Verification, validation, and evaluation in information security risk management. *IEEE Security & Privacy*, 9(2):58–65.

[Fielding et al., 1999] Fielding, R., Gettys, J., Mogul, J., Frystyk, H., Masinter, L., Leach, P., and Berners-Lee, T. (1999). RFC 2616, Hypertext Transfer Protocol – HTTP/1.1.

[Flood and Carson, 1993] Flood, R. L. and Carson, E. R. (1993). *Dealing with complexity: an introduction to the theory and application of systems science*. Plenum Pub Corp.

[Florencio and Herley, 2007] Florencio, D. and Herley, C. (2007). A large-scale study of web password habits. In *Proceedings of the 16th ACM International Conference on World Wide Web*, pages 657–666.

[Fodor and Roubens, 1994] Fodor, J.C. and Roubens, MR (1994). *Fuzzy preference modelling and multicriteria decision support*. Springer.

[Fowler, 2004] Fowler, M. (2004). *UML distilled: a brief guide to the standard object modeling language*. Addison-Wesley Professional.

[Freeman, T. and Housley, R. and Malpani, A. and Cooper, D. and Polk, W., 2007] Freeman, T. and Housley, R. and Malpani, A. and Cooper, D. and Polk, W. (2007). RFC 5055: Server-Based Certificate Validation Protocol (SCVP). *Internet Engineering Task Force*.

[Freier et al., 2011] Freier, A., Karlton, P., and P.Kocher (2011). RFC 6101: The Secure Sockets Layer (SSL) Protocol Version 3.0.

[Gambetta, 2000] Gambetta, D. (2000). Can we trust trust. *Trust: Making and breaking cooperative relations*, pages 213–237.

[Garlan and Shaw, 1994] Garlan, D. and Shaw, M. (1994). *An introduction to software architecture*. World Scientific Pub Co.

[Gaw and Felten, 2006] Gaw, S. and Felten, E.W. (2006). Password management strategies for online accounts. In *Proceedings of the 2nd ACM Symposium on Usable Privacy and Security*, pages 44–55.

[Gefen et al., 2003] Gefen, D., Srinivasan Rao, V., and Tractinsky, N. (2003). The conceptualization of trust, risk and their electronic commerce: the need for clarifications. In *Proceedings of the the 36th IEEE Annual International Conference on System Sciences, Hawaii* .

[Glade, 2012] Glade, B. (ed.) (2012). Identity Assurance Framework: Assurance Levels. Kantara Initiative.

[Glässer and Vajihollahi, 2010] Glässer, U. and Vajihollahi, M. (2010). Identity management architecture. *Security Informatics*, pages 97–116.

[Gnutella, 2003] Gnutella (2003). Protocol for a Revolution. `http://rfc-gnutella.sourceforge.net/`. [Online; accessed 06-February-2013].

[Gómez Mármol et al., 2010] Gómez Mármol, F., Girao, J., and Martínez Pérez, G. (2010). TRIMS, a privacy-aware trust and reputation model for identity management systems. *Computer Networks*, 54(16):2899–2912.

[Gómez Mármol and Martínez Pérez, 2010] Gómez Mármol, F. and Martínez Pérez, G. (2010). Towards pre-standardization of trust and reputation models for distributed and heterogeneous systems. *Computer Standards & Interfaces*, 32(4):185–196.

[Goodner and (eds.), 2009] Goodner, M. and (eds.), A. Nadalin (2009). Web Services Federation Language (WS-Federation) Version 1.2. OASIS Standard.

[Gopalakrishnan, 2009] Gopalakrishnan, A. (2009). Cloud computing identity management. *SETLabs briefings*, 7(7):45–54.

[Grandison and Sloman, 2000] Grandison, T. and Sloman, M. (2000). A survey of trust in internet applications. *Communications Surveys & Tutorials, IEEE*, 3(4):2–16.

[Grefen et al., 2009] Grefen, P., Eshuis, R., Mehandjiev, N., Kouvas, G., and Weichhart, G. (2009). Internet-based support for process-oriented instant virtual enterprises. *IEEE Internet Computing*, 13(6):65–73.

[Groß, 2003] Groß, T. (2003). Security analysis of the SAML single sign-on browser/artifact profile. In *19th Annual IEEE Computer Security Applications Conference*, pages 298–307.

[Guide, ISO, 2002] Guide, ISO (2002). 73: 2002. *Risk management–Vocabulary–Guidelines for use in standards.*

[Haenni, R., 2005] Haenni, R. (2005). Using probabilistic argumentation for key validation in public-key cryptography. *International Journal of Approximate Reasoning*, 38(3):355–376.

[Hammer-Lahav, 2010] Hammer-Lahav, E. (2010). RFC 5849: The OAuth 1.0 protocol. *Internet Engineering Task Force (IETF).*

[Han et al., 2010] Han, J., Mu, Y., Susilo, W., and Yan, J. (2010). A generic construction of dynamic single sign-on with strong security. *Security and Privacy in Communication Networks*, pages 181–198.

[Hardt et al., 2007] Hardt, D., Bufu, J., and Hoyt, J. (2007). OpenID Attribute Exchange 1.0 - Final. `http://openid.net/specs/openid-attribute-exchange-1_0.html`. [Online; accessed 06-February-2013].

[Hardt, 2005] Hardt, Dick (2005). Identity 2.0. *presentation at OSCON*, 2005.

[Hickman, 1995] Hickman, Kipp E.B. (1995). The SSL Protocol.

[Higgins Project, 2009] Higgins Project (2009). Higgins project website. `http://www.eclipse.org/higgins/`. [Online; accessed 06-February-2013].

[Hoffman et al., 2009] Hoffman, K., Zage, D., and Nita-Rotaru, C. (2009). A survey of attack and defense techniques for reputation systems. *ACM Computing Surveys (CSUR)*, 42(1):1.

[Howlett et al., 2012] Howlett, J., Hartman, S., Tschofenig, H., Lear, E., and Schaad, J. (2012). Application Bridging for Federated Access Beyond Web (ABFAB) Architecture, draft-ietf-abfab-arch-04 (Informational). Technical report, IETF Network Working Group.

[Hoyt et al., 2006] Hoyt, J., Daugherty, J., and Recordon, D. (2006). OpenID Simple Registration Extension 1.0. `http://openid.net/specs/openid-simple-registration-extension-1_0.html`. [Online; accessed 06-February-2013].

[Hughes et al., 2005] Hughes, J., Cantor, S., Hodges, J., Hirsch, F., Mishra, P., Philpott, R., and (eds.), E. Maler (2005). Profiles for the OASIS Security Assertion Markup Language (SAML) v2.0. OASIS Standard.

[Hughes and Maler, 2005] Hughes, J. and Maler, E. (2005). Security Assertion Markup Language (SAML) V2. 0 Technical Overview. *OASIS SSTC Working Draft, sstc-saml-tech-overview-2.0-draft-08.*

[Internet2, 2008] Internet2 (2008). Distributed Dynamic SAML. `https://spaces.internet2.edu/display/dsaml/Distributed+Dynamic+SAML`. [Online; accessed 06-February-2013].

[ITU, 2000] ITU (2000). X. 509: Information technology-open systems interconnection-the directory: Public-key and attribute certificate frameworks. *ITU-T Recommendation.*

[ITU-T Focus Group on Identity Management (FG IdM), 2007] ITU-T Focus Group on Identity Management (FG IdM) (2007). Report on Identity Management Use Cases and Gap Analysis.

[Jansen, 2010] Jansen, Wayne (2010). *Directions in security metrics research.* DIANE Publishing.

[Jaquith, 2007] Jaquith, A. (2007). *Security metrics: replacing fear, uncertainty, and doubt.* Addison-Wesley Professional.

[Jen and Lee, 2000] Jen, L. and Lee, Y. (2000). Working Group. IEEE recommended practice for architectural description of software-intensive systems. In *IEEE Architecture.*

[Jensen, 2012] Jensen, J. (2012). Federated Identity Management Challenges. In *7th IEEE International Conference on Availability, Reliability and Security (ARES)*, pages 230–235.

[Johansson, 2012] Johansson, L. (2012). RFC 6711: An IANA registry for Level of Assurance (LoA) Profiles. Technical report, IETF.

[Jones and (eds.), 2009] Jones, M.B. and (eds.), M. McIntosh (2009). Identity Metasystem Interoperability Version 1.0. OASIS Standard.

[Jøsang, 1996] Jøsang, A. (1996). The right type of trust for distributed systems. In *Proceedings of the 1996 ACM Workshop on New Security Paradigms*, pages 119–131.

[Jøsang et al., 2005] Jøsang, A., Fabre, J., Hay, B., Dalziel, J., and Pope, S. (2005). Trust requirements in identity management. In *Proceedings of the 2005 Australasian Workshop on Grid Computing and e-research* , ACSW Frontiers '05, pages 99–108, Darlinghurst, Australia. Australian Computer Society, Inc.

[Jøsang et al., 2007] Jøsang, A., Ismail, R., and Boyd, C. (2007). A survey of trust and reputation systems for online service provision. *Decision Support Systems*, 43(2):618–644.

[Jøsang and Presti, 2004] Jøsang, A. and Presti, S. (2004). Analysing the relationship between risk and trust. *Trust Management*, pages 135–145.

[Kamvar et al., 2003] Kamvar, S.D., Schlosser, M.T., and Garcia-Molina, H. (2003). The eigentrust algorithm for reputation management in p2p networks. In *Proceedings of the 12th ACM International Conference on World Wide Web*, pages 640–651.

[Keeney and Raiffa, 1993] Keeney, R.L. and Raiffa, H. (1993). *Decisions with multiple objectives: preferences and value trade-offs*. Cambridge University Press.

[Kellomaki and Wason, 2003] Kellomaki, S. and Wason, T. (eds.) (2003). Liberty ID-SIS Personal Profile Service Implementation Guidelines.

[Kemp et al., 2005] Kemp, J., Cantor, S., Mishra, P., Philpott, R., and (eds.), E. Maler (2005). Authentication Context for the OASIS Security Assertion Markup Language (SAML) v2.0. OASIS Standard.

[Kremer et al., 2002] Kremer, S., Markowitch, O., and Zhou, J. (2002). An intensive survey of fair non-repudiation protocols. *Computer Communications*, 25(17):1606–1621.

[Kylau et al., 2009] Kylau, U., Thomas, I., Menzel, M., and Meinel, C. (2009). Trust Requirements in Identity Federation Topologies. In *International Conference on Advanced Information Networking and Applications, 2009. AINA '09.*, pages 137 –145.

[Landau and Moore, 2011] Landau, S. and Moore, T. (2011). Economic Tussles in Federated Identity Management. In *Tenth Workshop on the Economics of Information Security (WEIS'11)*.

[lasso, 2010] lasso (2010). Liberty Alliance Single Sign-On. `http://lasso.entrouvert.org/`. [Online; accessed 06-February-2013].

[Liberty Alliance, 2013] Liberty Alliance (2013). Liberty Developer Tutorial. `http://www.projectliberty.org/liberty/content/download/423/2832/file/tutorialv2.pdf`. [Online; accessed 06-February-2013].

[Luhmann, 1979] Luhmann, N. (1979). *Trust and power*. Wiley Chichester.

[Luna et al., 2011] Luna, J., Ghani, H., Vateva, T., and Suri, N. (2011). Quantitative Assessment of Cloud Security Level Agreements: A Case Study. *Proceedings of the International Conference on Security and Cryptography (SECRYPT 2012).*

[Madsen, 2009] Madsen, P. (2009). Liberty IGF Privacy Constraints Specification, Version 1.0. Liberty Alliance Project.

[Maler et al., 2005] Maler, E., Cahill, C.P., Hughes, A.O.L.J., Origin, A., Beach, M., Metz, B.R., Hamilton, B.A., Randall, R., Wisniewski, T., and Reid, E.I. (2005). Security and Privacy Considerations for the OASIS Security Assertion Markup Language (SAML) V2. 0.

[Maler and Reed, 2008] Maler, E. and Reed, D. (2008). The venn of identity: Options and issues in federated identity management. *Security & Privacy, IEEE*, 6(2):16–23.

[Manchala, 2000] Manchala, D.W. (2000). E-commerce trust metrics and models. *IEEE Internet Computing*, 4(2):36–44.

[Marichal, 1998] Marichal, J.L. (1998). Aggregation operators for multicriteria decision aid. Doctoral dissertation, Institute of Mathematics, University of Liège, Liège, Belgium.

[Mármol and Pérez, 2009] Mármol, F.G. and Pérez, G.M. (2009). Security threats scenarios in trust and reputation models for distributed systems. *Computers & Security*, 28(7):545–556.

[Marsh, 1994] Marsh, S.P. (1994). *Formalising trust as a computational concept. Phd Thesis.* University of Stirling.

[McGarty, 1999] McGarty, C. (1999). *Categorization in social psychology.* SAGE Publications, London.

[McKnight and Chervany, 1996] McKnight, D.H. and Chervany, N.L. (1996). The meanings of trust.

[Mendoza et al., 2011] Mendoza, Florina Almenárez, Marín, Andrés, Sánchez, Daniel Díaz, Cortés, Alberto, Campo, Celeste, and García-Rubio, Carlos (2011). Trust management for multimedia p2p applications in autonomic networking. *Ad Hoc Networks*, 9(4):687–697.

[Ministerio de Administraciones Públicas de España, 1999] Ministerio de Administraciones Públicas de España (1999). *Magerit: Risk Analysis and Management Methodology for Information Systems.* Manuales (España. Ministerio para las Administraciones Públicas).: Serie Administración general.

[Monjas et al., 2009] Monjas, M.A., Yelmo, J.C, Trapero, R., and Bayona, R.M (2009). El broker de identidad como herramienta de gestión de la información de identidad. In *Telecom I+D.*

[Nadalin et al., 2006] Nadalin, A., Goodner, M., Gudgin, M., Barbir, A., and (eds.), H. Granqvist (2006). NIST SP 800-63 Electronic Authentication Guidance - NIST Special Publication 800-63, Version 1.0.2.

[Nadalin et al., 2009] Nadalin, A., Goodner, M., Gudgin, M., Barbir, A., and (eds.), H. Granqvist (2009). WS-Trust 1.3. OASIS Standard.

[Nadalin et al., 2004] Nadalin, A., Kalin, C., Hallam-Baker, P., and (eds.), R. Monzillo (2004). Web Services Security: SOAP Message Security 1.0.

[Nanda and Jones, 2008] Nanda, A. and Jones, M.B. (2008). Identity Selector Interoperability Profile v 1.5. Microsoft Corporation.

[OECD Report, 2011] OECD Report (2011). Digital Identity Management, Enabling Innovation and Trust in the Internet Economy.

[OIX, 2013] OIX (2013). Open Identity Exchange. `http://openidentityexchange.org/`. [Online; accessed 06-February-2013].

[OMNeT++, 2012] OMNeT++ (2012). The OMNeT++ Simulation Framework. `http://www.omnetpp.org/`. [Online; accessed 06-February-2013].

[OpenCA Research Labs, 2013] OpenCA Research Labs (2013). Libpki project. `http://www.openca.org/projects/libpki/`. [Online; accessed 06-February-2013].

[OpenID, 2007] OpenID (2007). OpenID Authentication 2.0 - Final. `http://openid.net/specs/openid-authentication-2_0.html`. [Online; accessed 06-February-2013].

[OpenLDAP, 2013] OpenLDAP (2013). The OpenLDAP website. `http://www.openldap.org/`. [Online; accessed 06-February-2013].

[OpenSSL, 2013] OpenSSL (2013). The OpenSSL website. http://www.openssl.org/. [Online; accessed 06-February-2013].

[OWASP, 2012] OWASP (2012). OWASP Risk Rating Methodology. https://www.owasp.org/index.php/OWASP_Risk_Rating_Methodology. [Online; accessed 06-February-2013].

[Pacyna et al., 2009] Pacyna, P., Rutkowski, A., Sarma, A., and Takahashi, K. (2009). Trusted identity for all: Toward interoperable trusted identity management systems. *IEEE Computer*, 42(5):30–32.

[Page et al., 1999] Page, L., Brin, S., Motwani, R., and Winograd, T. (1999). The pagerank citation ranking: bringing order to the web.

[Palomino and Rivero, 2008] Palomino, A.E. and Rivero, J.M.S. (2008). *Ohsas 18001: 2007*. FC Editorial.

[PICOS Project, 2011] PICOS Project (2011). Privacy and Identity Management for Community Services. http://www.picos-project.eu/. [Online; accessed 06-February-2013].

[PRIME Project, 2008] PRIME Project (2008). Privacy and Identity Management for Europe web site. https://www.prime-project.eu/. [Online; accessed 06-February-2013].

[PrimeLife Project, 2011] PrimeLife Project (2011). Privacy and Identity Management in Europe for Life. http://www.primelife.eu/. [Online; accessed 06-February-2013].

[Project, 2012] Project, Liberty Alliance (2012). Introduction to the Liberty Alliance identity architecture. Technical report. [Online; accessed 06-February-2013].

[Purdy, 2010] Purdy, G. (2010). ISO 31000: 2009-setting a new standard for risk management. *Risk analysis*, 30(6):881–886.

[Recordon and Fitzpatrick, 2006] Recordon, D. and Fitzpatrick, B. (2006). OpenID Authentication 1.1. http://openid.net/specs/openid-authentication-1_1.html. [Online; accessed 06-February-2013].

[Recordon et al., 2008] Recordon, D., Jones, M., Bufu, J., Daugherty, J., and Sakimura, N. (2008). OpenID Provider Authentication Policy Extension 1.0. http://openid.

net/specs/openid-provider-authentication-policy-extension-1_0.html. [Online; accessed 06-February-2013].

[Rescorla, 1999] Rescorla, E. (1999). RFC 2632: Diffie-Hellman Key Agreement Method.

[Rescorla, 2001] Rescorla, E. (2001). SSL and TLS: Designing and Building Secure Systems.

[Rescorla, 2005] Rescorla, E. (2005). RFC 4101: Writing Protocol Models.

[Resnick and Zeckhauser, 2002] Resnick, P. and Zeckhauser, R. (2002). Trust among strangers in Internet transactions: Empirical analysis of eBay's reputation system.

[Saaty, 1990] Saaty, T.L. (1990). How to make a decision: the analytic hierarchy process. *European journal of operational research*, 48(1):9–26.

[Sabater and Sierra, 2005] Sabater, J. and Sierra, C. (2005). Review on Computational Trust and Reputation Models. *Artificial Intelligence Review*, 24:33–60.

[Sanchez et al., 2012] Sanchez, R., Almenares, F., Arias, P., Díaz-Sánchez, D., and Marín, A. (2012). Enhancing privacy and dynamic federation in IdM for consumer cloud computing. *IEEE Transactions on Consumer Electronics*, 58(1):95 –103.

[Sánchez Guerrero et al., 2012] Sánchez Guerrero, R., Arias, P., Almenares Mendoza, F., and Díaz-Sánchez, D. (2012). Trust-aware federated IdM in consumer cloud computing. In *IEEE International Conference on Consumer Electronics (ICCE)*, pages 53 –54.

[Sengupta et al., 2011] Sengupta, S., Kaulgud, V., and Sharma, V.S. (2011). Cloud Computing Security–Trends and Research Directions. In *IEEE World Congress on Services (SERVICES)*, pages 524–531.

[Sheckler, 2007] Sheckler, V. (2007). Liberty Alliance Contractual Framework Outline for Circles of Trust.

[Singh, 2005] Singh, Ray P. (2005). ITU-T Focus Group on Identity Management (FG IdM): Report on IdM Use Cases and Gap Analysis. http://www.itu.int/dms_pub/itu-t/oth/15/04/T15040000040001PDFE.pdf. [Online; accessed 06-February-2013].

[Skogsrud et al., 2004] Skogsrud, H., Benatallah, B., Casati, F., and Dinh, M.Q. (2004). Trust-Serv: a lightweight trust negotiation service. In *Proceedings of the International Conference on Very Large Data Bases*, pages 1329–1332.

[SLAatSOI Project, 2011] SLAatSOI Project (2011). Empowering the Service Economy with SLA-aware Infrastructures. `http://www.sla-at-soi.eu/`. [Online; accessed 06-February-2013].

[Smart, 2010] Smart, N. (2010). ECRYPT II yearly report on algorithms and keysizes (2009-2010).

[Solberg, 2011] Solberg, A.Å. (2011). Interoperable SAML 2.0 Web Browser SSO Deployment Profile. 2011. `http://saml2int.org/profile/0.1`. [Online; accessed 06-February-2013].

[Solhaug et al., 2007] Solhaug, B., Elgesem, D., and Stølen, K. (2007). Why Trust is not proportional to Risk. In *Proceedings of the 2nd International Conference on Availability, Reliability and Security (ARES)*, pages 11–18.

[Somorovsky et al., 2012] Somorovsky, J., Mayer, A., Schwenk, J., Kampmann, M., and Jensen, M. (2012). On breaking saml: Be whoever you want to be. In *USENIX Security*.

[Squicciarini et al., 2008] Squicciarini, A.C., Czeskis, A., and Bhargav-Spantzel, A. (2008). Privacy policies compliance across digital identity management systems. In *Proceedings of the SIGSPATIAL ACM GIS 2008 International Workshop on Security and Privacy in GIS and LBS*, pages 72–81.

[Stoneburner et al., 2002] Stoneburner, G., Goguen, A., and Feringa, A. (2002). Risk management guide for information technology systems. *NIST Special Publication*, 800(30):800–30.

[Sun et al., 2010] Sun, S.T., Boshmaf, Y., Hawkey, K., and Beznosov, K. (2010). A billion keys, but few locks: the crisis of web single sign-on. In *Proceedings of the 2010 ACM Workshop on New Security Paradigms*, pages 61–72.

[Suriadi et al., 2009] Suriadi, Suriadi, Foo, Ernest, and Jøsang, Audun (2009). A user-centric federated single sign-on system. *Journal of Network and Computer Applications*, 32(2):388–401.

[SWIFT Project, 2010] SWIFT Project (2010). Secure Widespread Identities for Federated Telecommunications. `http://www.ist-swift.org/`. [Online; accessed 06-February-2013].

[SymLabs, 2012] SymLabs (2012). ZXID: Open SAML implementation in C. `http://www.zxid.org`. [Online; accessed 06-February-2013].

[TERENA, 2012] TERENA (2012). Terena TF-EMC2: REFEDs Federation Survey. `https://refeds.terena.org/index.php/Federations`. [Online; accessed 06-February-2013].

[The Internet Society, 2008] The Internet Society (2008). Trust and the Future of the Internet.

[Tran and Wietfeld, 2009] Tran, T. and Wietfeld, C. (2009). Approaches for optimizing the performance of a mobile SAML-based emergency response system. In *13th IEEE Enterprise Distributed Object Computing Conference Workshops (EDOCW 2009).*, pages 148–156.

[Triantaphyllou, 2000] Triantaphyllou, E. (2000). *Multi-criteria decision making methods: a comparative study*, volume 11. Kluwer Academic Publishers Dordrecht.

[Varney and Sheckler, 2005] Varney, C. and Sheckler, V. (eds.) (2005). Deployment Guidelines for Policy Decision Makers.

[Vaughn Jr et al., 2003] Vaughn Jr, R.B., Henning, R., and Siraj, A. (2003). Information assurance measures and metrics-state of practice and proposed taxonomy. In *Proceedings of the 36th Annual Hawaii International Conference on System Sciences.*

[Vervest and Zheng, 2009] Vervest, P.H.M. and Zheng, L. (2009). The network experienceŮnew value from smart business networks. *The Network Experience*, pages 3–14.

[Wallace et al., 1996] Wallace, D.R., Ippolito, L.M., and Cuthill, B.B. (1996). *Reference information for the software verification and validation Process*, volume 500. DIANE Publishing.

[World Wide Web Consortium (W3C), 2007] World Wide Web Consortium (W3C) (2007). Web Services Policy 1.5 - Framework. Technical report, http://www.w3.org/TR/2007/REC-ws-policy-20070904/.

[Xiang et al., 2010] Xiang, Y., Kennedy, J.A., Richter, H., and Egger, M. (2010). Network and Trust Model for Dynamic Federation. In *ADVCOMP 2010, The Fourth Interna-*

tional Conference on Advanced Engineering Computing and Applications in Sciences, pages 1–6.

[Zuo et al., 2010] Zuo, Y., Luo, X., and Zeng, F. (2010). Towards a dynamic federation framework based on SAML and automated trust negotiation. *Web Information Systems and Mining*, pages 254–262.