4-24-2015 12:00 AM

Privacy in Cooperative Distributed Systems: Modeling and Protection Framework

Afshan Samani
The University of Western Ontario

Supervisor
Hamada H. Ghenniwa
The University of Western Ontario

Graduate Program in Electrical and Computer Engineering
A thesis submitted in partial fulfillment of the requirements for the degree in Doctor of Philosophy

PRIVACY IN COOPERATIVE DISTRIBUTED SYSTEMS: MODELING AND PROECTION FRAMEWORK

Thesis Format: Monograph

by

Afshan Samani

Electrical and Computer Engineering

A thesis submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy

The School of Graduate and Postdoctoral Studies
The University of Western Ontario
London, Ontario, Canada

Abstract

A new form of computation is emerging rapidly with cloud computing, mobile computing, wearable computing and the Internet-of-Things. All can be characterized as a class of "Cooperative Distributed Systems" (CDS) in *open environment*. A major driver of the growth is the exponential adoption by people and organizations within all aspects of their day-to-day matters. In this context, users' requirements for privacy protection are becoming essential and complex beyond the traditional approaches. This requires a formal treatment of "privacy" as a fundamental computation concept in CDS paradigm.

The objective is to develop a comprehensive formal model for "privacy" as base to build a CDS based framework and platform in which various applications allow users to enjoy the comprehensive services in open environments while protecting their privacy seamlessly. To this end, this thesis presents a novel way of understudying, modeling and analyzing privacy concerns in CDS. A formal foundations and model of privacy is developed within the context of information management. This served as a base for developing a privacy protection management framework for CDS. It includes a privacy-aware agent model for CDS platform with the ability to support interaction-based privacy protection.

The feasibility of the proposed models has been demonstrated by developing an agent-based CDS platform using JIAC framework and a privacy-based Contract Net Protocol. It also included the application scenarios for the framework for privacy protection is Internet-of-Things, cloud-based resource scheduling and personal assistance within the project of smart space.

Keywords

Cooperative Distributed System (CDS), Interaction, Privacy, Sensitive Information, Computation.

Acknowledgments

I would like to express my special appreciation and thanks to my supervisor Professor Dr. Hamada H. Ghenniwa, you have been a tremendous mentor for me. I would like to thank you for encouraging my research and for allowing me to grow as a research scientist. Your advice and wisdom were always a support for my research and career as well as my life. Special thanks to examiner committee Dr. Ali Ghorbani, Dr. Michael Katchabaw, Dr. Weiming Shen and Dr. Luiz Capretz for making the defense a great experience for me with a lot of supports and insights to my work. I would also like to thank Dr. Abdolmotalib Wahaishi for his invaluable assistance in my research. I would specially thank my colleagues at CDS-Eng research group Dr. Raafat Aburukba, Ali Hussain, Shawn Talbot, Adrian Bienkowski for their support and encouragements.

A special thanks to my family. Words cannot express how grateful I am to my mother and father for their sacrifices and prayers that sustained me thus far. I would also like to thank my friends Dr. Razieh Samimi and Bahman Daei who supported me in writing, and incented me to strive towards my goal. At the end I would like express appreciation to my beloved husband Pouya for his patients and spending sleepless nights with me and always supporting me in the moments when there was no one to answer my queries.

Table of Contents

List of Figures

.

Symbols and Notations

The following is the list of symbols and notation frequently used in this work.

Notation/Symbol	Concept
W	World that is a CDS-based environment
e_i	A computation entity in CDS environment
	i: is the entity identity
I_i	Set of information that is owned by e_i
	i: is the entity identity
O_i	Set of operations that is owned by e_i
	i: is the entity identity
$E_{i,k}$	Exposure Boundary of $I_{i,k}$ that includes entities for which sharing $I_{i,k}$ can take place without causing privacy concern.
	i: is the entity identity
	k: is the information identifier
$I^S(I_{i,k}, e_j)$	$I_{i,k}$ is Sensitive in relation with e_j from e_i perspective
	i: is the entity identity that owns the information
	j: is the entity identifier that does not belong to $E_{i,k}$
	k: is the information identifier
$\bar{\bar{o}}(I^{x1}, I^{aux}, I^{x2})$	Executing Operation (o) on explicit information I^{x1} to transform the implicit information to explicit form of I^{x2}
$\bar{\bar{o}}(\widetilde{I^{x1}, I^{aux}})$	Preventing/Neutralizing Execution of operation (o) on I^{x1} given the auxiliary information I^{aux}
$S(I_{i,k}, e_j)$	Sharing $I_{i,k}$ with e_j
	i: is the entity identifier that owns the information
	j: is the entity identifier that receives $I_{i,k}$
	k: is the information identifier
$D(I_{i,k}, e_j)$	Disclosure of $I_{i,k}$ to e_j
	i: is the entity identifier that owns the information
	j: is the entity identifier that $I_{i,k}$ is disclosed to
	k: is the information identifier

$\hat{O}_j^{i,k}$	Non Authorized operations in O_j that can be applied on $I_{i,k}$ i: is the entity identifier that owns the information j: is the entity identifier that $I_{i,k}$ is disclosed to k: is the information identifier
$\hat{O}_j^{i}$	All possible non authorized operations in relation with e_j i: is the entity identifier that owns the information j: is the entity identifier that can receive information from e_i
$PV(e_j, I_{i,k}, \hat{O}_j^{i,k}, \theta_{i,j}^{i,k})$	Privacy Violation of e_i by e_j disobeying the agreement $\theta_{i,j}$ between e_i and e_j by executing a non-authorized operations belonging to $\hat{O}_j^{i,k}$ on $I_{i,k}$
$PP(e_j, (PS(I_i)), \hat{O}_j)$	Privacy protection of e_i when I_i is the space and $\hat{O}_j$ is all possible non authorized operations in e_j
μ	Privacy Protection Mechanism
$\bar{\bar{\mu}}$	Applying privacy protection mechanism
$PPL(e_j, I_i, \mu)$	PPL: probability of privacy protection of $e_i$using μ protection mechanism in interaction with e_j
IP	Interaction protocol
R^*	Participating Entities in an interaction protocol
I_i^S	All sensitive information in e_i in relationship with entities in R^*
S_M	Sequences of messages in an interaction protocol
$ss_{q,t}$	Sub-sequences of a sequence q: Sequence identifier t: sub-sequence identifier
$ss_{q,t}^{o}$	All operations of a sub-sequence q: Sequence identifier t: sub-sequence identifier
$\bar{\bar{ss}}_{q,t}^{o}(M)$	Execution of operations of a subsequence q: Sequence identifier t: sub-sequence identifier
$\mu_{i,k}$	Protection Operation in a computation entity that is applied for protecting $I_{i,k}$ that is classified as sensitive

Chapter 1

1 Introduction

Computation history is replete with changes in how people regard, use and interact with computers. With the recent evolution of computation from the colossal machines to the ever-present digital era that is characterized by technologies such as nanotechnologies, quantum computing, cloud-based computing, mobile computing and the new area of computation known as Internet-of-Things (IoT), a great paradigm shift through which many technological services have become part of nearly every human activity. In spite of the beneficial comforts that are experienced due to these technological services, the use of information technology reveals the extent to which there could be a risk to privacy.

1.1 Cooperative Distributed System and Privacy Concerns

In an increasingly interconnected, intricate and quickly changing world, more entities choose to connect and do business online. Both people and businesses are engaging with various applications and because of this, it is envisioned that a significant part of our lives will be steered by computation systems in near future. It is estimated that in 2020, there will be 6.58 smart internet-connected devices for each person [1]. Despite the development of computation environments in delivering services to people and businesses, privacy is still a major challenge in these environments [2], [3].

The evolution of Cooperative Distributed Systems (CDS) created new forms of computation that instituted the significant advances, involvement and tremendous impacts of information technology on peoples' lives. In CDS, autonomous self-interested entities require the capabilities of others, resulting in interaction and exchange of information between these entities. It is envisioned that information is collected by many processes and devices and hence has brought increased risks regarding the concerns on one's privacy. Information about people is gathered through many service providers, stored in various infrastructures, analyzed and reported for further objectives [4]. The information is manipulated towards extracting and disseminating the information to other parties or serving various interests. For example, smart house applications capture sensor

information from separate parts of the building. Some of these sensors might collect the electrical consumption rate of each apartment. The smart house manager application may collect this information and send it to the power company. The aggregated view of this information from all apartments in question can be used for designing and distributing power plans in neighborhoods. However, the individualized type of this information can reveal the time that householders are not available at their home based on identifying the pattern of lowest consumption rate at each apartment. This suggests that disseminating some information such as the individualized view of the power consumption can lead to additional information about them. This also illustrates that the exchange of information among entities in CDS environments can cause a level of concern from these entities for their privacy. In particular, in open CDS environments, it would be a strong assumption that entities in the environment will have a degree of respect for the privacy of others.

The computation in distributed heterogeneous environments that are modeled as CDS occurs during interaction between entities where the information is *shared*. This entails capturing privacy at the computation level [5]. This view is contrary to the traditional approaches towards privacy through which the application filters the computation solutions based on predefined rules [6], [7]. The privacy models can be classified into two main categories: rule-based approaches and architectural-based approaches [8]. Privacy solution models that evolve from rule-based approaches are typically designed for stable, low variant environments. These approaches mainly concentrate on applying rules onto information that is collected during the process of *sharing*. Due to the open environment assumption in many applications of CDS, the rule-based approaches [9] are not sufficient [8], [10]. Information processing has been the engine of extracting information by applying operations on it. This information is not necessarily captured in rule-based privacy models. Furthermore, since the rules and policies can impose limitation of the design and dynamism of the environments, many open CDS environments cannot adopt these perspectives on privacy.

Among architectural-based privacy solutions are anonymization techniques [11], [12], [13], privacy utility trade off mechanisms, [5], [14], [15] social tradeoffs and proxy-based privacy protection [16]. In this context, the anonymization techniques are limited to

particular settings that include a trusted information collector entity and non-continuous information dissemination processes for which it cannot be adopted by open CDS environments [17]. The work in [16] illustrates that privacy utility trade off models do not necessarily reflect the preferences that each entity might have over their privacy. The utility tradeoff mechanisms have been applied in contexts such as smart power grid in which privacy is reduced to limited access to individualized signal from the aggregated view of the collected signal [15]. These models also evolved with approaches for measuring the risk of privacy concerns. Such risk adheres to the execution of operations that causes privacy concern but it can measure the probability of the entity's information being used [18]. In all cases, the limitation of the proposed models indicates the lack of adequate privacy model for CDS.

It is noteworthy that privacy is correlated with the interaction aspects of computation systems. This asserts that privacy is a computation concept that is related to the interaction process and can be adequately addressed by interaction protocols. For instance, if a specific entity e_i can reach solution S_1 by acquiring the capabilities of entity e_j, the devised interaction protocol for such engagement has to coordinate the pertinent activities with e_j. However, during this engagement, e_j may exploit the information as part of the messages in the interaction protocol and thus could result in privacy concern for e_i. Capturing privacy as a concept in interactions still adheres to the mechanism of interaction as well as finding solutions that may not be conducive to privacy concerns for the participant entities.

1.2 Privacy: Concepts, Issues and Models

Privacy is an ethical, a social and a legal concept that has gained substantial definitions. The Merriam-Webster Dictionary defines privacy as "the state of being alone: the state of being away from other people" while the Oxford Dictionary defines it as "the state in which one is not observed or disturbed by other people". In all definitions, privacy becomes an inherent feature of an environment of multiple people (entities/agents) or a setting of decentralized entities/agents. Decentralized computation environments can be adequately molded as CDS [19].

In an information management model of computation, “privacy” contains some specific connotations though in many ways the term is similar to how it is generally understood. In communication-based interaction among entities, it becomes a privacy concern when *sensitive* information flows outside the entity or the unit of entities in CDS. Evidently, it will be a more difficult challenge in CDS in particular when communication-based interactions are applied in open environments.

Motivated by the computational view on privacy, understanding privacy concept that can be applied in contexts such as CDS requires formal analysis of privacy. The work in [20] proposes a formal approach for capturing privacy in information management in the context of social networks. However, the analysis stays at formulating the norms and relationship of the roles, and the concept of privacy is not clearly stated. In addition, the concept of norms and contexts can be implicit and exist in gray areas when it comes to social networks [21]. In another work in [22], an extensive grammar and syntax of a language is provided for expressing the privacy policies enforced by HIPPA through which privacy becomes limited to policy context that is applied. Different approaches and many privacy models have been proposed to deal with relevant privacy issues [23], [7], [24]. However; to our knowledge, none of these approaches have treated and captured privacy at the computational level adequate for the CDS environments.

There have been significant efforts towards building a foundation for privacy rights during digital interactions. This enables an understanding of privacy and adopting the associated concepts based on practices in information technology law [25], [26]. Many countries have enacted laws and legislations to protect people’s privacy. For instance, the Canadian law has several legal acts that oblige service providers and consumers to be responsible on respecting privacy as a right for people. Canadian Information Privacy Act and Access to information are among these legal supports. Furthermore, some privacy models were motivated by the supporting legal scenarios and rules [22]. Due to limitations on the setting of the rules and scenarios, employing these models impose *closed* assumption on the environment.

1.3 Scope of the thesis

A major objective of this work is to conduct a deep analysis of "privacy" and to develop a formal model and computation concepts of privacy concerns. Also, it attempts to utilize the formal model to develop a privacy protection frame work for CDS-based applications.

In many cases privacy studied and treated in conjunction or within the context of "security" and "trust". Although practically these concepts might be directly related, within this thesis, however, our focus was on analyzing the foundation of privacy and developing a fundamental model as computation concept in CDS paradigm. Our belief is privacy is an intrinsic concept. In this work, privacy is viewed within the context of managing information manipulation, in particular "sensitive" information, within a given exposure boundary, for a given of security and trust measurements. Where, "security" mechanisms concern about the truthfulness of the communication within the areas of confidentiality, integration and availability. And "trust" is defined as the degree of belief of reliability among entities in a particular context. This direction makes the principle foundations of our findings expandable to model and address situations where security and trust are involved.

1.3.1 Formal Privacy Model

The lack of a formal privacy model that is applicable for a CDS was the motivation to develop a formal treatment of privacy in CDS environments. The proposed model is used as an analytical tool to evaluate the state of the privacy during any entity's interaction.

Entities discern the sensitivity of information differently depending on the recipients of the information in an interaction. Sensitive information perceived by one entity might be considered totally as a non-sensitive in relation to another. Entities tend to not *share* information, when it is labeled as sensitive. This creates an exposure boundary for entities' information which positions privacy as the state of the exposure boundary of the information. Information within the exposure boundary is non-sensitive but becomes sensitive when it exists outside of the exposure boundary.

Information exists in explicit form. However, it can be classified as implicit information when it is in conjunction with operations. Operations can retrieve explicit information by processing the said information. The execution of operations transforms the implicit information to explicit form. Through this, information might be transferred to outside of the exposure boundary and become sensitive. This implies that the concern with privacy is about the *disclosure* of sensitive implicit information. For example, various IaaS (Infrastructure as a Service) [27] providers serve their consumers by offering them resources, including memory, storage, and computational power, among others. In many forms of IaaS service delivery models, payment packages (pay per user) are based on the demand of entities. When it is not serving a higher priority consumer, economical packages receive response from the server. The advantage of costly packages is the guarantee of service at any time. Hence, serving an economical plan at the server implicitly implies not having a high priority job. *Sharing* scheduling information may enable an entity with medium priority and a resource-demanding job to acquire the service provider. Frequent preemption for lower priority consumers may lead to service blocking. This explains that *sharing* the schedule is *not sensitive* when in possession of the scheduler, but it is *sensitive* to *share* with other consumer entities.

In this work, we have provided an original privacy model that formally captures the concepts and concerns about privacy. Within this model, privacy concerns, privacy violation and privacy protection are formally explained and the necessary concepts to develop a framework for privacy protection management are introduced.

1.3.2 Privacy Protection Management Framework

By employing the proposed privacy model, we established a privacy protection management framework that incorporates privacy protection mechanisms at the interaction level. Because achieving perfect privacy protection requires complete knowledge about the world, we proposed quasi protection mechanism that can protect privacy with a certain level of probability that is addressed as Privacy Protection Level (PPL).

The framework captures the information of entities and accordingly evaluates the exposure boundaries associated to information. Consequently, it identifies the sensitive information and determines the necessary extension form for privacy protection. Using the PPL measure of each mechanism, the PPL of the privacy-based interaction protocol is evaluated and this enables applications to adopt privacy mechanisms that generate an acceptable level of PPL at the interaction level. It is formally proven in this work that the protection can sufficiently occur at the interaction level and the privacy-based interaction protocol has quantifiable PPL.

1.3.3 Privacy-Aware Computation Platform

In order to capture privacy at the computation platform, it has to be treated as a mathematical object. The computation system reduces the available solution choices to the ones that can fulfill the expected privacy requirements. The quantifiable model for privacy concept allows filtering the solutions space based on privacy measures as well as maximizing the privacy protection in interactions among entities. For example, scheduling solutions collect scheduling variables and boundaries to reach to global schedule for the participant entities. Typically, privacy concerns are not incorporated as the scheduling criteria for which the schedule might not be acceptable. The computation view on privacy enables scheduler to capture privacy as solution boundaries or decision variables that results in scheduling solutions for which privacy is protected. This example will be discussed in more details in Chapter 7.

The proposed privacy protection framework can be applied as an analytical tool to evaluate the state of privacy in interactions of entities as well as being applied at contexts such as computation. The computation entity employs the privacy protection management to extent the interaction protocol to a privacy-based interaction protocol through which the solution inherit privacy at the computation. We have formally proven that the resolution to privacy is part of the computation solution. In this work, we have extended the computation entity in an agent-based model [19] by introducing the privacy protection management and implemented the privacy-aware entity using the Java-based Intelligent Agent Componentware (JIAC) platform [28].

Privacy is an immense area of research that has attracted many researchers, scientists and developers within the computer science and engineering arenas. The tremendous work devoted on the perspectives of the authorization and rule management within underlying infrastructure [27], [6], privacy related concepts and the challenge with the new technologies [22], taxonomy of privacy affairs [29], [10], privacy categorization and personally identifiable information [8], privacy within the context of information management including information collection, information processing and information dissemination [23, 30], [31]. There also have been some attempts in formalizing the languages used for privacy policies [22]. The economic mechanisms have been applied in this area as well with the objective of developing strategies through which privacy protection would be a dominant strategy [32]. Furthermore, privacy has been the concern of multi-agent systems. Agents interact on behalf of their principals, engage in a number of activities and exchange information, which inevitably raises issues and concerns with regard to privacy[16]. Our research has contributed in several aspects of these areas, which is *shared* with privacy in information management, formalizing privacy concepts, personally identifiable information, privacy concepts and categorization and privacy within multi-agent systems.

A major contribution of this work is to develop a privacy-aware computation in open Cooperative Distributed Systems that addresses and manages privacy at the interaction level and thus provides a certain degree of privacy protection at the interactions of entities. The work introduces several new original and novel ideas that contribute to the overall thesis that can be listed as follows:

1) The formal modeling of privacy in the context of information management

Formal analysis of privacy concepts is essential in capturing privacy as a computation concept. In this work, we have investigated privacy within the context of information management and sensitive information. Our attempts in understanding privacy in this context results in developing a formal model that delivers a complete view on privacy in information management.

2) An Interaction-based Privacy Protection Management Framework.

Considering the incomplete knowledge of entities in open CDS environments, privacy protection is encountered with an uncertainty level. To deal with the uncertainty, a probability-based model is applied. The privacy protection framework enables managing the expected level of privacy protection within the interactions of entities. The proposed solution for protecting privacy has been congregated within an architectural approach towards interaction-based framework for privacy protection in which the privacy protection mechanisms are applied to interactions as it is required.

3) Expressing privacy as a computation concept.

The privacy concept is formally treated at the computation level by including privacy in the computation solution. As a result, the computation entity adopts the privacy protection management as part of the computation entity architecture.

4) A Privacy-Based Interaction Protocol

Applying privacy protection management framework on interaction protocols allows identifying privacy concerns at the interactions. It evaluates the messages and sequences of the interaction protocol and provides adequate protection operations within the interaction protocol that result in privacy-based interaction protocol. The extended privacy based interaction protocol that is generated by applying the privacy protection management framework can sufficiently provide privacy protection in situations where the knowledge in CDS environment is incomplete. One of the interaction protocols that are utilized within this framework is Contract Net protocol (CNP). CNP is a negotiation based interaction protocol that is designed for distributed problem solving. Due to privacy concerns in this protocol, we have applied the privacy protection management framework that resulted in a privacy-based Contract Net interaction protocol.

5) A Quantifiable Privacy Protection Level for the privacy based interaction protocol.

With the proposed approach in the privacy protection management framework, the protection level of the mechanisms can be measured and can hence provide quantified

values. As a result, the privacy based interaction protocol is able to define the level of privacy protection that the protocol provides.

6) Applying the privacy aware computation entity in a Service Oriented Semantic Driven Architecture (SOSDA) Environment.

The proposed privacy-aware computation entity can be integrated with CSD-ENG smart space applying SOSDA principals [33] where the interaction protocol is providing the privacy protection. The implementation challenges of expanding the Collaborative Intelligent Rational (CIR) agent architecture to include privacy solutions as part of the computation solution are elaborated and resolved in Java-based Intelligent Agent Component Ware (JIAC) [28].

1.4 Organization of Thesis

The rest of this work is organized as follows: Chapter 2 provides an overview on privacy in different areas of research. Chapter 3 provides a novel approach for a formal modeling of privacy. Subsequently, Chapter 4 proposes a privacy protection management framework. Chapter 5 elaborates on privacy-aware computation platform by expanding the computation entity and its implementation challenges. Chapter 6 presents the application of the privacy protection management framework on Contract Net protocol. The applicability of the proposed privacy protection management framework in various different environments and application domains is outlined in Chapter 7. The future work and the conclusion of this work is presented in Chapter 8.

Chapter 2

2 Background and Literature Review

Despite the comfort that is experienced with new information technologies, they have imposed privacy concerns on people and businesses. The more people engage with digital developments, the more are concerns for their privacy. Primarily, privacy concerns were studied and practiced in legal communities and researches. However, privacy has become inseparable challenge of nowadays' digital interactions in which it carries tremendous amount of information about people. Many disciplines have addressed privacy in their solutions however, adequate privacy models for CDS environments is still a challenge.

2.1 Privacy In law[1]

Privacy is a multi-disciplinary concept that is mainly tented within Law researches and legal schemes. Understanding privacy from the perspective of law enables us to observe and perceive privacy concerns in the context of information management. There are various views about privacy among different categories of law. One believes privacy is the product of the modern life where gossips became curiosity while another claims that privacy is as old as common law [25]. The work in [36] indicates that privacy is often interpreted as security and it is traded in return for providing security for the society or individual [26]. The concept of privacy has been studied in four main categories [25]:

- Common Law
- Constitutional Law
- Statutory Law
- International Law

[1] The term "Privacy Violation" has been used here to reflect the concept addressed in the law. This term will be formally defined later in the next chapters.

Due to dynamic context of privacy, challenge in front of legal scholars is defining privacy rights which, in many cases are typically abstract and vague [25]. Researchers in legal areas try to retrieve the potentials of the existing law to propose solutions for protecting privacy and evaluate Law responses to new subjects such as privacy rights. Traditionally, privacy was treated as "decisional privacy" which mainly concerns the liberty of decisions about one's body and family. Nonetheless, because of the role of technology in spreading information about people and organizations and the direct effect of privacy in ones' lives, it has become the priority in legislative agenda in Congresses. History of privacy rights indicates multiple stories about people and organizations in which dissemination of information can directly target individuals' lives [25].

One of the main achievements in Privacy Law is presenting it as one's "Rights". The main issue in the current technology is the presence of medias that are utilized for circulating information. Such trend increases the effect of privacy in people's lives. Therefore, attorneys typically address privacy rights in the area of "common law". The objective is to protect privacy of private lives from unwanted intrusion. Accordingly, there are four type of intrusion in interaction of people and society [25]:

1. Intrusion upon seclusion and solitude.
2. Public disclosure of embarrassing private facts.
3. Publicity which exposes people in a false light in public.
4. Appropriation for people's interests.

As people's lives are now virtually available among various type of services and data sources, it would become essential for these services to adapt their solution in alignment with common law. However, privacy rights are not limited to common law and people's private life. More importantly, privacy concerns are not only about people. It can also be applied on how machines and software interact which can be addressed in information privacy. In this section, we try to extract the necessary foundation for privacy interactions so that we can associate them in general interaction among entities in CDS.

In attempt to identify the interactions that result in privacy violation from law perspective, four types of violation categories are presented above. Each of which can represent various circumstances that individuals or machines confront in open environments. For instance, the first category asserts on respecting people's solitude and private avocations. This implies that the actions performed by an entity in its private life are being monitored by another entity apart from their awareness. This is equivalent to the privacy concerns related to "information collection" and "information processing". Currently, digital life is an inseparable part of individuals' activities [25]. However, mainly, all the individual's online private affairs and activities are usually monitored and recorded by service providers. Software and machines are installed in many locations to observe and analyze human interactions. The motivations supporting these systems are tailored to improving business, security, better consumer support, safety, efficiency and many human perspectives. Yet, such motivations has brought about and created a tremendous challenge related to privacy in Cyberspace. Nonetheless, legal efforts are directed to finding solutions that can mitigate the issue by eliminating unnecessary monitoring and controlling tasks. The second Category implies the concern of public exposure of information, which might cause humiliation and embarrassments for individuals [25]. This is due to the *sharing* an individual's information to others without having the necessary consent. This form of privacy concerns is referred as secondary use whenever a third party is involved. With the explosion of Internet Media and personal pages in various web sites, individuals experience levels of disconcertion when their information is used in other contexts. Personal information is excessively spreading among Internet services and in noticeable amount of cases; it has been disseminated to other providers or publicly exposed.

Similar to the second category, the third category of intrusion occurs when disclosing false information entails the attraction of unnecessary attention to individuals [25]. Suppose in a reputation system built for auctions, an entity gets false negative feedback; it is without doubt that such falsification impact further future activities with this entity. Spreading false information about capabilities and availability of a service provider in a grid environment can forge the scheduling mechanism and hence may overload a provider or disrupt the whole scheduling system.

The last category of intrusion discusses the appropriation of exposing individuals' interest information [25]. Due to the possibility of extracting personal information about people by processing their interests in various subjects, interest information become sensitive. Given the growth of targeting advertisement, interest information is valuable to advertisers. This could exhibit levels of privacy concerns when the interest information is not appropriate.

As argued in [25], the challenge in investigating privacy violation is distinguishing the discussed aforementioned categories. For simplicity, they are addresses respectively as 1) intrusion, 2) disclosure, 3) false light and 4) appropriation. In spite of the similarity among these categories, they have characteristics that assist in separating the concepts. For instance, in intrusion and disclosure, existence of secret information is part of the scenario. In disclosure and false light, the publicity is the main element. However, in false light, falsified information or fiction differentiates it from disclosure. Appropriation typically involves in providing advantages for the owner of information [25].

Borrowing the intrusion categories in common low, similar concerns exist in cyber space. Among them are: "Breach of Confidentiality", "Defamation", "Infliction of emotional distress" and "privacy of home" [25].

Breach of Confidentiality": this term commonly is used to define the revealing of patients' and client's information [25]. In this context, the patient is the consumer entity and the doctor is the service provider. If the service provider breaches the confidentiality of the information, it has disseminated the information to a third party without having the consent of the consumer.

Defamation refers to disrupting individuals' reputation by false information [25], where Infliction of emotional distress is related to the emotional discomfort that individuals experience when their sensitive information is *shared* in social networks and similar communication mediums.

The Privacy of home concept addresses the physical resident of individuals. This is associated with ones' solitude and private affair that are well established in common law.

This type of privacy concern can infiltrate to individuals' digital interactions when their information is spread across various sectors in machine.

2.2 Privacy in Information Management

Privacy has been defined from the perspective of multiple views. For example privacy has been interpreted as: "the freedom of thought", "having control over one's body", "the solitude in one's home, "the freedom from surveillances", "the protection of one's reputation", "protecting one from searches and interrogation" and "not selling one's information" [29]. Also, there are fundamental legal privacy theories such as: Privacy is the limited access to self [34], privacy is the right to be left alone [35], secrecy in many legal communities were accepted as definition for privacy [36], Control over personal information [37], Intimacy and Personhood [38] also were numerated as theories of privacy. In addition, the work in [10] addressed privacy as "the condition of being protected from unwanted access by others" [38]. Similarly, the work in [9] defines the privacy as the right to determine "to what extent information about people or companies is communicated to others". Adopting similar concepts in the context of information, privacy can be adequately treated in "information management" in CDS environments. Entities in CDS autonomously interact and *share* information through which it can be processed or disseminated. Due to self-interestedness and autonomy of the entities in CDS settings, there might be privacy concern at each of the levels of information management.

Information management can be categorized based on the nature of actions or operations applied including [29]:

- Information Collection: the process of compiling information such as surveillance, online profiling, online tracking, collecting task specification and requirements.

- Information processing: applying operations on information such as aggregation, integration and identification.

- Information dissemination: the process of publishing or diffusing information to other entities such as breach of confidentiality and increased accessibility.

2.2.1 Personally Identifiable Information (PII)

Information or attributes such as SIN numbers and personal number can be used to identify entities. Some attributes can be used in combination of others to identify an entity; for example, combination of date of birth, gender, name and zip code. The attributes that directly identify the entities are called "identified" and the attributes that can [implicitly] result in identifying an entity are called "Personally Identifiable Information" (PII). In this context, attribute *disclosure* happens when the value of identifiable information reveals the identity of the entity. And, identity *disclosure* happens when the identifiable information is a bridge to associate sensitive attributes to an entity[39]. The challenge is that due to advances in technology and information processing which can convert the non-PII attributes to PII attributes at higher scale, it becomes not possible to directly identify PII[8].

Entities' incomplete knowledge in open environments originates the concern on the operations that might be applied on *shared* information. Combining information by applying operations to extract new information is known as a *secondary use problem*. This could lead to privacy concerns when the retrieved information is sensitive and the information includes the identifier to the owner of the sensitive information. This issue which is functionally equivalent to the PII problem is due to *implicitly* extracting information from identifiable information that is *shared* [40],[10]. Resolving the PII problem has been investigated in three approaches; reduction, expansion and PII2.0. **Reduction** focuses more on "identified" attributes. For example, COPPA (Children Online Privacy Protection Act) concerns only with information about "identified person". In fact, the "identifiable" concept has been reduced from this approach. In the **Expansion** approach, the identifiable information is considered as critical as identified information. However, as almost any kind of information can be attributed to an identified entity, and

from the practicality point of view, this approach is considered as a flaw. This is the result of treating the identified and identifiable information equally[40][2].

PII 2.0 is an approach for privacy in interactions that deals with PII problem through the perspective of risk analysis. Although, there are large amount of identifiable information, that could implicitly retrieve new identified information, not all of them have a high risk of privacy concerns. PII 2.0 introduces the risk of revealing information as a relative probability measure. If the risk of a set of identifiable information is high, then information should not be *shared* [40]. The risk of interaction is probabilistic view of the occurrence of associated negative impact of privacy concerns on the entity. It allows decision-making processes to evaluate the interaction and the *sharing* information with regards to the risk of interaction, gain and the possible drawback that might affect the entity.

In new forms of resolutions for PII complications, there are rule-based and standard-based approaches. Typically, the rule-based approaches are convenient when the area of social and technological development have reached a fairly stable state [13]. Due to the dynamic and open nature of environments in CDS, the rule-based solutions to resolve PII are not adequate approaches.

2.2.2 Human Everyday Privacy Model

In every day interactions, humans follow a conceptual privacy model that is affected by perception of humans on internal and external factors. Internal factors include [41]:

- Information Sensitivity (IS): entity's judgment on the sensitivity of the information
- Information Receiver (IR): entity's evaluation on the level of "trust" to the recipients of the information

[2] The concepts related to identified and identifiable information are formally treated as explicit and implicit information that are discussed in more details in Chapter 3.

- Information Usage (IU): entity's assessment over the gain of using information or the cost of the mistreatment of the information.

Additionally, there are external factors such as Laws (L), Market (M), Norms (N), and Architecture (A) [technological context of communication] of the interaction, Contextual variables (C) [set of traditional contextual variables such as activity, location, companions,.., etc.] that can impact the perception of humans over their privacy state. Therefore, the everyday conceptual privacy can be modeled as[41]:

$$Preferred\_privacy\_level = user(L, M, N, A, C, I, IS, IR, IU)$$

Where I is the *shared* information. The combination of L, M, N, A, C, I, IS, IR, IU can be interpreted as "situation". This reduced the model to [1]:

$$Preferred\_Privacy\_Level = user(situation)$$

Due to similarity of preferences of humans over different situations, the work in [1] captures the similarities as "Face".

$$Face = user(situation)$$

The "faces" will become different in various contexts. Also, the variables affecting the "situation" might be numerous [1]. This results in difficulties and inconveniences in capturing and applying the "Face" model in CDS environments where entities have distinctive interdependencies.

2.3 Privacy in Distributed Systems

Within the arena of distributed systems, privacy is a concern when the setting of the environment is decentralized. Distributed Systems can be classified in more granular categories that we address a few of them and discussed the related privacy models.

2.3.1 Privacy in Authorization Framework

Security and privacy in many cases have been interchangeably used where privacy is treated in the context of security. Traditionally, in the context of information

management, privacy was investigated at security authorization mechanisms[9], [42]. Despite security mechanisms that are targeted to maintaining confidentiality, integrity and availability of the communication among entities, privacy concerns are about manipulating the information that could have been securely communicated [*shared*]. The efforts within security mechanisms are geared towards assuring the information is to be only accessible by the desired entity, and the entities' communication is not compromised with a third party. However, security mechanisms may not address the manipulation of information among entities. For instance, the communication with a search engine can have the required security measures and the integrity of the communication is supported. Nonetheless, the information that is retrieved by the search engine after applying operations on the collected information is not treated in security mechanisms. This indicates that the nature of security mechanisms is not sufficient to resolve privacy concerns. Privacy concerns are categorized on the control over *"how"* information is collected, processed and disseminated. Typically, the security mechanisms are applied on the established connection between at least two entities. If the confidentiality, integrity and availability of the communicated information are satisfied, that interaction is secured. Nevertheless, that does not guarantee that there is not privacy concerns with the interaction.

Diverse set of models has been applied on authorization in CDS such as SAML, Akenti, PERMIS, Shibboleth, VOMS, XACML, GT4 [9] and [42]. The objective of these models is to provide authorization platforms that protect information from unauthorized access. However, these models are still incapable of addressing privacy in relation with *"how"* information is processed and "*flow"* within entities. Additionally, the solutions do not provide privacy protection techniques for the collection and the dissemination of information. The work in[42] addresses privacy as part of the populated rules for the authorization mechanism. However, the model does not capture the identifiable information that implicitly can lead to privacy concerns. In addition, the setting of the applied model in this mechanism is assumed to include trusted entities to govern the privacy rules. Such setting is not necessarily attainable in all CDS environments and the privacy model cannot be applied.

2.3.2 Privacy in Multiple Data Sources

Data source providers provide aggregated view of the information that is collected from people, business, and organizations. Typically, this information is published for research collaboration purposes and data analysis for a particular problem. However, the process of information collection can be pursued if exclusively, the aggregated information is published. *Disclosing* information such as the participation of an entity in the information collection process can lead to privacy concern for the entity. Many public data sources contain information that might be common across multiple data sources. Linking the available information across multiple data sources is based on their common information can identify individuals and *disclose sensitive information* which can be captured as identity *disclosure* and attribute *disclosure* [43], [12]. These concepts depend on contextual variables, amount of released data, level of the knowledge of adversary[39], [13]. Given this categorization, there are different privacy models that address specific aspects of privacy. Models such as K-Anonymity [11], l-Diversity [12], SIPPA [13], t-closeness [44] and Differential Privacy [38] aim to resolve identity or attribute *disclosure*. The typical setting of anonymization mechanisms includes a trusted information collector that collects the information and disseminates aggregated information to other entities [23], [44]. There are assumptions in this setting that the information collector is a trusted party and the process of information collection and dissemination happens in non-continuous fashion [45]. These mechanisms are tailored towards protecting sensitive information such as participation of entities in information collecting process. The adversary consumes the aggregated information in conjunction with previous knowledge to retrieve sensitive information about an entity. Evidently, not all CDS applications can adhere to the setting of anonymization mechanism. Furthermore, because of possibilities of attacks such as complementary attack in K-Anonymity[43], these approaches are not applicable in CDS. In complementary attack, the adversary accesses the published anonymized information in multiple sources and combines them all. This in many cases circumvents the protection that is applied.

2.3.3 Privacy in Distributed Constraint Satisfaction

Distributed Constraint Satisfaction Problem (DisCSP) is a Constraint Satisfaction Problem (CSP) in which the variables and constraints are distributed among distributed multiple entities (i.e., Agents). Those agents need to determine values for a set of variables such that the cost of a set of constraints over the variables is sataisfied and thus optimized (as either minimized or maximized). In other words, CSP is about finding a consistent assignment of values to variables[46, 47]. The DisCSP framework was a focal point of several areas such as Artificial Intelligent and agent Technology. In DisCSP, privacy principles have been identified at four level [47]namely: 1) The Agent, 2) The Topology, 3) The Constraint and 4) The Decision. At the Agent level, the algorithm has to guarantee that no agent can learn the identity of any other agent unless they are in *sharing* coordination constraints. At the topology level the algorithm should not allow any agent to learn about the constraints and cycles of other agents. For example, the constraint of an agent for specific resource is sensitive information that should be kept private. The Constraint level is similar to topology level with focus on constraint and its relations. Finally at the decision level, the algorithm has to protect the outcome of any decision that the agent makes. The solution in [47] expands the Distributed Pseudotree Optimization Procedure (DPOP) algorithm [48] by adding privacy metrics. This algorithm creates a Depth First Search Tree (DFS tree) out of entities. Each entity interacts only with their neighbors. Entities send their constraint to their parent, and the root node (leader) accordingly solves the problem and sends it back to others. The contribution of the solution in [47] anonymizes the construction of DFS. Nodes have code names for interactions. Moreover, the leader in each round is anonymous and given the associated assumptions, the approach can guarantee the required privacy levels. However, the settings in these environments are limited to the topology that is defined in priori and the maximum distance between two nodes in the environment which is known for the used algorithm. Evidently, the adoption of the solutions in DisCSP in CDS will not inherent to all settings of application. Furthermore, in this algorithm, it is possible for a malicious entity to forge the coordination information in attempt to be the leader which may perform actions that can cause privacy concern.

In addition, there are attempts to resolve privacy concerns in DCOP (Distributed Constraint Optimization Problem) [49], [50]. DCOP consists of entities that set and control the evaluation of variables. Entities decide which evaluation of the variables has more benefit for them. However, the problem's setting is based on the assumption that all entities are aware of the constraints of other entities, and only the evaluation of the variables is sensitive information [50]. Additionally, privacy solutions in DCOP are derived from an information theoretic perspective [50] and do not necessarily reflect on the privacy concern in setting in CDS environment

2.4 Privacy in Distributed Artificial Intelligence

Multi Agent System (MAS) is one of the computational models applied in CDS in which the computational entities operate in a decentralized control fashion and modeled as autonomous entities known as agents. MASs are designed for autonomous actions and flexible interaction [51]. Agents act on behalf their principals and engage in various interactions that might require in many cases the exchange of personal information[16]. This, as such makes privacy management an essential aspect.

Privacy management approaches in MASs have been categorized into three categories: (i) policy-based, (ii) privacy utility tradeoff and (iii) social relationships. For instance, the work in [30] is a policy-based framework in which a trusted broker compares the policies of providers and consumers and decides on their compatibility. The broker resumes any interaction only if the compared policies are compatible. However, the approach relies on the assumption that the broker is a trusted entity[16]. The Privacy Enhancement Agent (PEA)[52] is a similar approach that uses P3P (Platform for Privacy Protection)[53]retrieve the P3P policies, validate the compatibility of policies and accordingly decide on the possibility of further engagement in any interaction.

Other approaches adopt the ontological comparison of policies that are described and represented using the Web Ontology Language (OWL) [54]. Once the conditions are accepted among both parties, the consumer *shares* the information. In similar approaches, the rules are semantically analyzed and the access control mechanism are incorporated

with the privacy rules [6][7]. However, in these models, there is a lack of mechanisms which obliged entities to comply with the commitments [16].

One of the major challenges in privacy management is to identify and measure the *risk* of *sharing* the information. To deal with such issue, "Privacy- Utility Tradeoff"' mechanisms were proposed[16],[5]. This work is based on calculating the information gain of *shared* information. The elements such as history of two sides of interaction, social aspects of interaction, relevancy of requested information to the offered service has not been considered in these mechanisms. This motivated the complementary approaches that applying concepts of trust and intimacy in measuring risk and utility. The challenge with these approaches is the difficulty of validating these metrics, in particular in CDS environments [16]. The utility trade off mechanisms evolved with approach of measuring the risk of privacy concerns. The risk of interaction adheres to execution of operations that might cause privacy concern but it can measure the probability of the entity's data getting used [18].

2.5 Privacy In Cooperative Distributed Systems

Many solutions are proposed for computations for which the environment is modeled as CDS. Typically, the prospects of these models are tailored towards particular setting of the environment where a certain type of information is exchanged in the interaction of entities. Adopting these solutions for many applications of CDS imposes limitations and assumption of their environments. In the following we address some of the related works within this area.

2.5.1 Privacy in Auction Mechanisms

Auctions are subclass of markets that restrict the governing rules of the market in which buyers and seller are trading goods and services. Auction mechanism design is the attempt to manipulate the rules of the auction in order to achieve specific goals[55]. In auction configurations, an auctioneer applies the rules of the auction mechanism and rewards the winner(s). In this setting, it is possible that a faulty or malicious auctioneer forges the auction or exploits the bidding values[56] When bidders submit their bids to

the auctioneer, it is possible that the auctioneer exploits the bidding value of the winner for the future auctions. For example, if the winner's bid is $900 and the second bid value is $600, then the auctioneer can start the auction from $900 since it has the knowledge that at least one entity will bid with this value [56]. It is very desirable and an important aspect of bidding activities to assure the bidders about the safety of the auction with respect to privacy concerns.

To deal with this issue some approaches were proposed in the literature [56] ,[57]. The work in[56] an Auction Issuer (AI) is introduced which is a passive entity that has no direct communication with bidders and limits the auctioneer ability to only access the relevant information. The AI in this architecture computes the auction and presents it back to the auctioneer. This restricts the auctioneer to be able only to know the identity of the winners only and not the value of the bids. However, this protocol cannot guarantee the privacy of entities when collusion takes place between the AI and auctioneer. The (AI) entity is designed to control the access of auctioneer entity to sensitive information.

2.5.2 Risk Analysis

Risk analysis in interactions of entities has played a significant role in many privacy solutions. Identifying risk levels in a system provides meaningful measures which can be applied to processes that could mitigate the risk [4]. Risk in general is a degree of belief on occurrence of an event with undesired outcomes. The risk of interaction refers to level of belief on incidents and events in which *sharing* information in interaction led to privacy concerns. There are various models to capture the risk of interactions. Some of them adhere to analyzing the interactions in terms of 1) Information Sensitivity, 2) Information Receiver, 3) Information Usage[41] Other approaches use fuzzy logic to capture the effecting variables on risk of interactions. The work in [58] utilizes hierarchical fuzzy inference system to address the risk of interaction. It measures and evaluates the relevancy of the requested information; trust level, cost and criticality of the shared information, type of intended operation, the content of the agreement, sensitivity of information and information gain in a given interaction. Using these variables, a hierarchical fuzzy system can be developed to measure the risk of interaction.

2.5.3 Targeting Advertisement

Targeting advertisement systems apply Online Behavioral Advertisement (OBA) techniques to promote more relevant commercial contents to users. Because of capability interdependency among entities of these systems, they need to exchange information such as user's interest that might be sensitive. In this context, privacy becomes a major challenge [59] [60]. One of the approaches in addressing privacy concerns is through Adnostic [59] In Adnostic system, privacy is modeled as a tuple that is expressed in terms of the following attributes <consumer's identity, consumer's request>. The disclosure of any relevant attribute may result in privacy concern to consumers. In this system, it was presumed that, providers are able of delivering their capability without knowing the identity of consumers. The objective of the model is to protect consumer's privacy by introducing a trusted entity called Trusted Third Party, (TTP). Providers and consumers are defined as roles, which can be played interchangeably. A provider has to present a list of options to the consumer whose in turn consumer selects the preferred information which will be considered as the request information. However, consumers encrypt the list of options including the one that was tagged as the chosen option. When providers receive the encrypted list, they only know that an item is selected but they are now aware which one is chosen [59]. In Adnostic, it is assumed that there is a time period where providers have to wait before providing their capabilities. In this time, they need to collect all encrypted lists of options sent by consumers, aggregate all these lists and submit them back to the TTP at the end of waiting period. The TTP is capable of decrypting the list and thus delivers the decrypted list to the provider. The provider's access to an aggregated list of requests does not show which identity has chosen which item in the list. Another approach in targeting advertisement is through decoupling the request and identity utilizing ElGamal crypto systems [60]. However, in these approaches, the protection mechanism can be circumvented if entities collide [40]. Furthermore, the only sensitive information in this model is the combination of consumer's identity and their requests. This makes the system incapable of managing various settings in CDS environments.

2.6 Summary

Despite the variety of works carried out toward protecting privacy in different disciplines, an adequate privacy model for CDS environment is lacking. Within the context of information management, privacy can be categorized as information collection, information processing, information dissemination and invasion. One of the challenges of the privacy concept is the identification, which is referred to manipulating information in order to retrieve and relate "sensitive information" to entities. However, information may have different risks for the identification. Identified information can directly lead to the risk of inferring and identifying an entity. The setting of these two categories is different, which makes it not possible to differentiate among them.

The Law perspective on privacy and information technology provides classifications on scenarios that can be realized in digital interactions. These scenarios are instances of the privacy concerns that could happen in CDS environments. Furthermore, privacy has been evaluated in many research subjects such as authorization mechanisms, publishing data sources, Multi Agent Systems, Distributed Constraint Satisfaction Problem, auction mechanisms, risk assessment and targeting advertisements. The authorization mechanisms tend to address access control issues, while attending to how information is used and manipulated is neglected. Most of privacy models in publishing data sources address the issues of publishing aggregated information by a trusted entity. This setting is not necessarily applicable in all applications of CDS environments. Solutions in MAS need to make entities comply with what they commit. They also need to consider the social aspects of the relationships between entities. Realizing such setting in CDS may not be feasible. Evaluating risks of interactions to address privacy was pursued in different privacy models. Nonetheless, to our knowledge, none of the existing privacy solutions have provided a privacy model that is adequate for CDS environments.

Chapter 3

3 Privacy Concerns in CDS: Concepts and Model

Privacy is the interest of immense area of research for which many models have been proposed that are automated in different applications. Some of the applied models require settings where impose limitations on the design of the entities of the environment and create a *closed* environment. This necessitates employing privacy models that can capture privacy as a computational concept for which formal analysis of privacy becomes essential. Treating privacy in information management context enables modeling privacy in a computation context where the flow of sensitive information becomes a concern for privacy. Our contribution in this chapter includes the formal analysis of privacy and modeling it in the context of information management.

3.1 CDS: Description and Agent-based Model

CDS are a class of systems in which entities are autonomous, self-interested, able to operate on some functions locally, and exercise some authority in *sharing* their capabilities. Goals in these settings refer to a state in which the actions of the entity, including physical and mental reasoning, are directed at the said state. Within CDS, entities have interdependencies through which some goals might be unattainable through the abilities of an individual entity. They may require coordinating activities with other entities to reach to an individual or collective goal state [19], [51]. This coordination is a class of solutions that provides structure and mechanisms to the system to deal with interdependency problems. "Structure" refers to the entities' pattern of communication and decision-making related to coordination. "Mechanisms" are a composition of decision points, coordinated control and interaction devices directed to resolve problems with interdependencies [19]. An essential characteristic of CDS is the distribution of control; this means that the strategies of entities cannot be controlled by outside parties. This supports the fact that every entity in CDS has a part of the solution in which participating entities' goals are achieved.

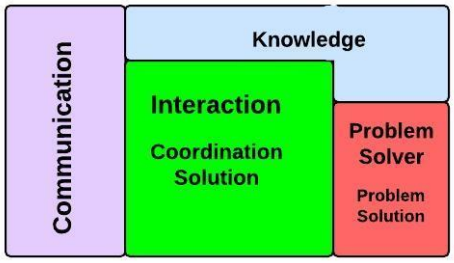

Figure 1. Computation Entity in CDS

In this work, we concentrate on entities of CDS in an Agent-Based model in which entities can be modeled as CIR agents. These Agents are organized by knowledge, problem solving, interaction, and communication capabilities [19]. "Knowledge" is the entity's mental state about the world, which is incomplete in many examples of CDS environment and the global knowledge is distributed among all entities. "Problem solving" refers to the entity's ability to identify the class of their goals, categorizing sub goals, applying required actions to the goals' state, and determining the type of interdependency. "Interaction" is the authority and capability of the entity in the pursuit of mechanisms that can resolve interdependency problems. Interaction mechanisms are steered by protocols that manage engagement between entities. The "communication" layer is responsible for packaging and transferring messages in the desired languages. [19] Communication-based interaction, or message-based Interaction, is essential when the entities' knowledge is incomplete and they are obligated to exchange messages. The connections between various aspects in computation entities are shown in Figure 1. Due to interdependency problems regarding settings in CDS, reaching a solution requires the interaction of multiple autonomous entities. This indicates that computation in CDS takes place within interactions among entities.

In the open structure of CDS environments, entities' availability and participation is unpredictable and there is no control on their behavior or the design that they adopt. The new form of computation emerging in Grid, cloud, and mobile computing can be modeled as open CDS. Cloud paradigms such as IaaS, PaaS, and SaaS have served many application domains such as medical, health, financial, entertainments, education, business, and communication.

3.2 Privacy Model and Analysis

Privacy is the concern of environments with multiple autonomous entities. Autonomy of entities facilitates them to exploit the information they receive in various ways. Consequently, this might reflect privacy of other entities. Considering a single entity in an environment, there will not be any privacy challenges. It technically is the natural characteristic of the environments that autonomous entities exchange information. Let W be the decentralized environment of autonomous self-interested entities.

$$W = \{e_1, \dots, e_N\}$$

In the context of information management, entities include information and operations. In the lowest granularity level, an entity can be shown as:

$$e_i = < O_i, I_i >, 1 \leq i \leq N, i: entities\ identifier$$

Where:

$$I_i \equiv \{I_{i,1}, \dots, I_{i,k}, \dots, I_{i,M}\},$$
$$1 \leq i \leq N, 1 \leq k \leq M, i: \text{entities}' \text{ identifier}, k: \text{information identifier}$$

and

$$O_i \equiv \{o_{i,1}, \dots, o_{i,w}, \dots, o_{i,W}\},$$
$$1 \leq i \leq N, 1 \leq w \leq W, i: \text{entities}' \text{identifier}, w: \text{operation identifier}$$

Privacy is in direct relation with self-realization of an entity, which reflects the objectives of the entity in a given environment [29]. In any situation that an entity has to be protected with respect to privacy, there is sensitive information that the entity doesn't prefer reveal and expose. Entities have various states. As an example, in an object oriented modeling, the entity might have different attributes in which each represents a state of an entity. Information is a tool for modeling the state of an entity. Therefore, the entity in relation with others desires to protect the sensitive states from being exposed to the outside world. Privacy also can be defined at the level of units. Family is the example of units in societies in which people have distinctive approaches in the flow of

information to the outside of the unit. In result, for any given state of an entity, there is a boundary for exposure. This suggests that Privacy is the state of exposure boundary of an entity's state with the outside world. For any information, there exists an exposure boundary including the entities that are considered to be inside the boundary. This is denoted as $E_{i,k}$ as the following:

$$E_{i,k} = \{e_{t+1}, \dots, e_{t+r}\}, \subset (E_{i,k}, W), 1 \leq r, t \leq N$$

Existing of boundary emphasizes on the flow of information within the boundary and preventing it from outside. This can describe the sensitive information as well. Information might be sensitive in relation with a particular entity and it might be non-sensitive with others. When the information flows within the boundary it is non-sensitive but existing of the information outside of exposure boundary makes it sensitive information. For instance, salary information is not sensitive to be *shared* with members of a family but it is when *sharing* it with a colleague. The exposure boundary is designated by the entity. Therefore, sensitive information is a relative concept reflecting the reluctance of an entity in *sharing* information with a particular entity.

$$I^S(I_{i,k}, e_j) \equiv \notin (e_j, E_{i,k}) \qquad 1.$$

Information exists in explicit form. It is the definite form of information. As previously described, information can be classified as sensitive and non-sensitive in relation with the entity being interacted with. It can also be classified as implicit information in relation to the operations that can be applied on the explicit information. Implicit information can be transformed to explicit information by execution of the operation.

The operation can be modeled as a function that extracts implicit information from explicit information and consequently transforms it to an explicit form. It also can combine the explicit information with other auxiliary information (denoted as I^{aux}) to transform the implicit information to explicit. The auxiliary information is collected or inferred information, which does not reflect any information by its own but it can expose information about an entity if it is used in combination with other information. This information can be an empty set of information as well, when the operation only needs

the given explicit information. Therefore, any implicit information is equivalent to some explicit information that can be defined as follows:

$$o(I^{x1}, I^{aux}) \equiv I^{x2} \quad 2.$$

Manipulation of explicit information by applying of operations can transform the implicit information into explicit form. In the above example,

$$\bar{\bar{o}}(I^{x1}, I^{aux}, I^{x2}) \quad 3.$$

illustrates that Executing Operation (o) on explicit information I^{x1} to transform the implicit information to explicit form of I^{x2}. In contrary

$$\bar{\bar{o}}(\widetilde{I^{x1}, I^{aux}}) \quad 4.$$

is used to show the execution of an operation is prevented or neutralized. Thus, the application of the operation cannot proceed.

One of the means of the flow of the information is through *sharing* information with other entities. Entities can decide if a particular entity belongs to the exposure boundary of certain information. When the entity is outside of the boundary, the information is considered sensitive and hence entities do not *share*. Therefore *sharing* is defined as a process that takes place only within the exposure boundary and can be formally expressed as:

$$S(I_{i,k}, e_j) \equiv \neg I^S(I_{i,k}, e_j) \wedge \left[= \left(I_j, \cup \left(I_{i,k}, I_j\right)\right)\right] \quad 5.$$

Although entities has the authority on protecting their relevant explicit sensitive information by not *sharing* it with others outside of the boundary, it becomes a concern when the implicit information might be transformed into explicit sensitive information. For instance, John's salary is classified as sensitive information. For example, John *shares* with Amy information, which states that his salary is 10 percent more than the average salary of the employees. If Amy has an operation that is capable of retrieving the employees average salary, she will be able to extract John's salary. In this example, the

statement “John’s salary is 10 percent above the employee’s salary” is explicit information, while Amy’s operations and this information implicitly refer to John’s salary which is considered being sensitive. This illustrates how implicit information may convey sensitive information and by transform it into explicit information will reveal the explicit sensitive information.

The presence of operations on the entity that receives the information results in possessing the implicit information.

$$D(I_{i,k}, e_j) \equiv$$

$$\left[= \left(I_j\ , (\cup\ (I_{I,k}, \bigcup_{w=1}^{W} (\ I_j, o_{j,w}(I_{I,k})), I_j\ ))\right)\right]$$

6.

By *sharing* non-sensitive explicit information, it is possible to *disclose* implicit information. The implicit information can be labeled as being sensitive or non-sensitive. This indicates that the *disclosure* of information might result in transferring the information to the outside of its exposure boundary. In other words, *privacy concern* relates to *disclosure* of *sensitive implicit* information.

Security mechanisms can provide the necessary control on the sharing process in which is applied at the exposure boundary. However, because the disclosure of sensitive implicit information can transfer the information outside of the exposure boundary, security mechanisms are not sufficient for managing privacy concerns.

In the previous example, $I_{john,k}$ is representing the statement “Jon’s salary is 10 percent employee average salary”. Amy also belongs to $E_{john,k}$ where implies $\neg I^S(I_{john,k}, e_{Amy})$. If Amy has a retrieval operation ($o_{Amy,ret}$) on a statistical dataset that includes the employee average salary I^{aux} and calculates john’s salary, $o_{Amy,ret}(I_{john,k}, I^{aux})$ is the implicit information that reflects John’s salary ($I_{john,k'}$).

$$\bar{\bar{o}}_{Amy,ret}(I_{john,k}, I^{aux}, I_{john,k'})$$

This suggests that if Amy executes $o_{Amy,ret}(I_{john,k}, I^{aux})$, she can extract John's salary. Disseminating information ultimately can be modeled by operations where the functionality of the operation is to transfer the information to other entities. As an example, Amy may perform an operation to send $I_{john,k}$ to Adam.

In another setting, let e_i and e_j be the entities interacting. Originally e_i has $I_{i,info1}$ and $I_{i,2}$ where $I_{i,info2}$ is sensitive to *share* and $I_{i,info1}$ is non-sensitive.

$$I^S(I_{i,info2}, e_j)$$

$$\neg I^S(I_{i,info1}, e_j)$$

However, e_j has operations in which can extract $I_{i,info2}$ from $I_{i,info1}$.

$$o(I_{i,info1}, I^{aux}) \equiv I_{i,info2}$$

$$\bar{\bar{o}}(I_{i,info1}, I^{aux}, I_{i,info2})$$

One of the main challenges of privacy relates to execution of operations that converts the sensitive implicit information to explicit form. Hence, having the knowledge about operations of the entity that receives the information can indicate what sensitive information can be retrieved by *sharing* of particular information. This introduces the concept of authorized operations. $O_j^{i,k}$ is a set of operations belonging to O_j where e_i has agreed on their application on $I_{i,k}$.

3.2.1 Privacy Concepts

Modeling privacy as a computational concept requires identifying measures that can reflect privacy in a computational model. In this section, we elaborate on the concepts that explain the state of privacy among interacting entities. These concepts have been applied for managing measures that can be associated to computational concepts.

When entities *share* information, they agree on the terms of utilization of the *shared* information. As an example, this can be enforced through the norms of various cultures in

people societies [25] or electronic legal agreements among web services[61] Ideally, these agreements include the allowed set of operations that can be applied on the *shared* information. Dishonoring a given agreement by the execution of non-authorized operations $\widehat{O}_j^{i,k}$ is considered to be an evidence of privacy violation. For instance in the above example, if e_j execute a non-authorized operation o , then it is said that e_j has violated the privacy of e_i . Similarly, if Amy applies the retrieval operation when it is not in the agreement with John, Amy has violated John's privacy. Accordingly:

$$-(\widehat{O}_j^{i,k}, O_j, O_j^{i,k})$$

where:

$$-(\theta, M, X) \equiv \forall\, x\ \ni (x, \theta) \mid \ni (x, M)\ \wedge \notin (x, X)$$

$$\widehat{O}_j^{i,k} \equiv \{\hat{o}_{j,1}^{i,k}, \dots, \hat{o}_{j,t}^{i,k}, \dots, \hat{o}_{j,T}^{i,k}\}, 1 \leq t \leq T$$

The unauthorized operations also can be defined in relation with all of information about an entity.

$$\widehat{O}_j^i = \bigcup_{k=1}^{M} (\emptyset, \widehat{O}_j^{i,k})$$

Based on the scope of communicated information through *sharing* and *disclosure*, the unauthorized operations can be applied on a subset of information (S) as well.

$$\widehat{O}_j^i(S) = \bigcup_{\forall s, \in (s, PS(S))} (\emptyset, \widehat{O}_j^s)$$

Non-authorized characteristics of an operation relates to the interacting entity. They can agree on the set of un-allowed. Let $\downarrow \widehat{O}_{j,w}^{i,k}$ the notation to address the negative permission over execution of an operation tagged as non-authorized. Entities agree on set of operations that cannot be executed over the *shared* information. This is considered to be the agreement between entity e_i and e_j while *sharing* $I_{i,k}$.

$$\theta_{i,j}^{i,k} \equiv \forall w \mid \downarrow \hat{o}_{j,w}^{i,k}$$

Given this, privacy violation of e_i by e_j is through disobeying the agreement $\theta_{i,j}$ between e_i and e_j by executing non authorized operations $o_{j,w}$ on $I_{i,k}$:

$$PV\left(e_j, I_{i,k}, \widehat{O}_j^{i,k}, \theta_{i,j}^{i,k}\right) \equiv \exists w \mid \theta_{i,j}^{i,k} \wedge [\bar{\bar{\hat{o}}}_{j,w}^{i,k}(I_{i,k})]$$

7.

While the privacy violation is about disobeying the agreement among entities, privacy protection is enforcing mechanism that prevents application of non-authorized operations on entities information. In the proposed model, any sets of information also are considered as information. Hence, the privacy protection is about preventing execution of non-authorized operations on all subsets of information.

$$PP\left(e_j, \left(PS(I_i)\right), \widehat{O}_j\right) \equiv \forall t, w \mid \subset (t, PS(I_i)) \wedge \widetilde{\bar{\bar{\hat{o}}}_{j,w}^{t}(t)} \quad 8.$$

$$PS(S) \equiv S' \mid \forall p, \in (p, S') \wedge \subset (p, S) \wedge (\nexists p' \mid \subset (p', S) \wedge \not\subset (p', S'))$$

Sharing information in CDS happens among entities during the interaction. In many cases, a privacy concern is geared to negative impacts on the owner of information. In this work, we have modeled the negative impact that might be resulted by privacy concern as the cost of interaction. It could be modeled by the negative utility that an entity perceives by exploitation of the information. For example, within the healthcare domain, various sensors and devices that are typically planted or embedded in the patient's body can provide different types of patient's information. If the gathered information is disclosed to the public because of a *sharing* process, a high risk of loosing job opportunities or insurance plans might be envisaged. Such a scenario reflects the cost of exploiting and *sharing* the sensors' information. Cost can be modeled by the perception of an entity about the negative impact that will be imposed by exploiting the *shared* information with a specific recipient.

$$C(I_{i,k}, e_j) \rightarrow R^{+}$$

However, in the context of incomplete knowledge, the model can be extended to capturing the risk of occurrence of such event. In the above example, the cost of losing insurance because of *sharing* medical information might be significant for many patients. However, such incidents rarely happen and the risk of it might not be high.

$$R(I_{i,k}, e_j) = P\left(\bar{\bar{\delta}}^{i,k}_{j,w}(I_{i,k}, I^{aux})\right) \times C(I_{i,k}, e_j))$$

3.2.2 Differential Privacy In Privacy Information Management Model

In this section, we position our findings in relation to Differential Privacy. We reduce the proposed formal privacy model to this model and illustrate that concepts of the model can be mapped to the proposed one.

Differential privacy is a model for creating randomized function that has been applied in various statistical databases including anonymized datasets. The setting in differential privacy includes an info collector e_i that provides aggregated information by gathering information from individuals. There are participants e_k that provide their information to the info collector and expect their information to stay private. Also, there are adversaries e_j that apply some operations on previous explicit information and others received to extend implicit sensitive information about participants. The objective of this model is to reduce the risk of *disclosing* individuals' information as the result of their participation in information collection process [23].

Info collector provides a set of operations [in the form of queries (Q)] that can be executed upon various entities including adversaries. An adversary already possesses auxiliary explicit information that the info collector is not aware of and utilization of this information results in privacy concern [23].

Modeling the setting of differential privacy using our proposed framework is as the following:

The world has at least three entities representing info collector, adversary and participant:

$$W = \{e_i, e_j, e_k\}$$

$$e_i =< I_i, O_i >$$

$$e_j =< I_j, O_j >$$

$$e_k =< I_k, O_k >$$

Participants *share* some information with info collector which is sensitive to *share* with another entity.

$$\mathrm{S}(\mathrm{I}_{k,\mathrm{l}}, \mathrm{e}_i) \rightarrow \neg \mathrm{I}^{\mathrm{S}}(\mathrm{I}_{k,\mathrm{l}}, \mathrm{e}_i)$$

$$\neg \mathrm{S}(\mathrm{I}_{k,\mathrm{l}}, \mathrm{e}_j) \rightarrow \mathrm{I}^{\mathrm{S}}(\mathrm{I}_{k,\mathrm{l}}, \mathrm{e}_j)$$

There is some auxiliary information about participants that is possessed by the adversary. It can be explicitly received or implicitly inferred.

$$\mathrm{D}(\mathrm{I}_{\mathrm{k,p}}, \mathrm{e}_{\mathrm{j}}) \rightarrow \in (\mathrm{I}_{\mathrm{k,p}}, I_j)$$

Equivalently, the information set of each entity is as the following:

$$\mathrm{I}_{\mathrm{i}} = \{\mathrm{D}, \mathrm{DB}, \mathrm{I}_{\mathrm{k,l}}\}$$

$$\mathrm{I}_{\mathrm{k}} = \{\mathrm{I}_{\mathrm{k,l}}, \mathrm{I}_{\mathrm{k,p}}\}$$

$$\mathrm{I}_{\mathrm{j}} = \{\mathrm{I}_{\mathrm{k,p}}\}$$

Assuming, there are queries $\mathrm{o}_{\mathrm{i,m}}$ at info collector e_{i} that can transform $\mathrm{I}_{\mathrm{k,l}}$ into new explicit information such as $\mathrm{I}_{\mathrm{i,b}}$. Similarly, there is an operation at adversary that utilizes $\mathrm{I}_{\mathrm{i,b}}$ to regenerate $\mathrm{I}_{\mathrm{k,p}}$:

$$\in (\mathrm{o}_{\mathrm{i,m}}, \mathrm{O}_{\mathrm{i}})$$

$$\in (o_{j,n}, O_j)$$

$$o_{i,m}(\{I_{k,l},\ D,\ DB\}) \equiv I_{i,b}$$

$$o_{j,n}(\{I_{k,p}, I_{i,b}\})\ \equiv\ I_{k,l}$$

Then application of $\hat{o}_{j,n}^{k,p}$ is not authorized and neglecting it inherently becomes the evidence of privacy concern.

$$\bar{\bar{\delta}}_{j,n}^{k,p}(\{I_{k,p}, I_{i,b}\})$$

The info collector applies a mechanism [differential privacy] to prevent the execution of $o_{j,n}$. Differential privacy mechanism enables the info collector to include noise information to the result of each query. The outcome is new information that cannot be used for retrieving $I_{k,l}$.

$$\text{PP: } o_{i,m'} | \bar{\bar{o}}_{i,m'}(\{I_{k,l},\ D,\ DB, I_{i,b}\}) = I_{i,b'} \wedge \bar{\bar{\delta}}_{j,n}^{k,p}(\{I_{k,p}, I_{i,b'}\})! = I_{k,l}$$

In above, we demonstrated the setting of differential privacy as privacy protection model using the proposed framework. In this section, we elaborate on concepts of differential privacy model and present a comparison between the proposed model and differential privacy.

Differential privacy is a quasi-protection mechanism which has been developed using the preventive manipulative approach in the context of statistical databases. This model has defined privacy as a goal to reducing the risk of entity being denied in a situation as the result of participating in a statistical database [23]. Based on this, privacy violation equivalently has been addressed by privacy breach concept. This concepts is the state of a Turing machine c in which is not halted if the adversary finds the correct s in a given database DB and its distribution D:

$$\text{Privacy Breach} \equiv \text{the adversary generates } S \text{ where } C(D, DB, S) \text{accepts}$$

The adversary is an entity that makes efforts in generating proper S to make C not halted. This definition implies that privacy concern is a function of the "explicit information" learned from the database.

Utilizing the differentially private randomizing functions is motivated by modeling privacy protection at participation of entities. In the other word, privacy protection is the state of producing outputs [explicit information] in which participation of any single entity does not impact the result in a huge extend. This argues that "participation of an entity in a statistical database" is the information that privacy protection is targeting. This suggests that "participation" is considered to be sensitive information.

Sensitivity is in direct relation with the perception of an entity about the recipients of information [41]. However, the above analysis illustrates that there is an assumption in differential privacy in which only considers the "ownership" of information as sensitive information. This is the reason that they capture sensitivity at the operation level. The result of all operations will be incorporated with levels of noise which can satisfy the conditions of differentially private functions. This has been captured using $L1-\text{Sensitivity}$ measure. For given datasets D_1 and D_2 differing only in one element and a query function [operation] f, $L1-\text{Sensitivity}$ has been defined as:

$$\Delta f = \max_{D_1,D_2}(||f(D_1) - f(D_2)||_1)$$

The mechanism in differential privacy adds noises to the result of queries to protect privacy of participants. The variance of the added noise is denoted as σ. To realize the conditions of $\epsilon-\text{Differential Privacy}$, σ has to be greater than $\frac{\epsilon}{\Delta f}$. All queries received by the info collector are examined through the above condition and the necessary noise is added to the result [62].

The above assumption limits the capabilities of the model to process more complicated scenarios. As an example, John e_{John} is a patient that suffers from a severe disease and he is under trial of a new research to find a cure for the disease. Let the disease be Ψ. John's medical information has been *shared* with a medical statistical database e_{sdb} to assist researchers with finding a cure. Also, some information such as the region Υ John comes

from is collected. This helps researchers to perform history analysis on their patients. After running experiments $o_{sdb,cure}$ on patients, they realize that there is 90 percent correlation between high cholesterol level K and affected by Ψ. As information is *shared* with the statistical database, they are not considered to be sensitive. However, when an insurance company e_{ins} investigates various patients, some information becomes sensitive.

In information management, sets and subsets of information are considered to be information as well. Therefore, in relation with e_{sdb} :

$$\neg I^S(I_{John,k}, e_{sdb}),$$

$$\neg I^S(I_{John,\Psi}, e_{sdb}),$$

$$\neg I^S(I_{John,Y}, e_{sdb}),$$

$$I_{John,O\Psi} = \{Owner(e_{John}, I_{John,\Psi})\}$$

$$\neg I^S(I_{John,O\Psi}, e_{sdb}),$$

$$I_{John,OK} = \{Owner(e_{John}, I_{John,k})\}$$

$$\neg I^S(I_{John,Ok}, e_{sdb})$$

$$I_{John,\Psi K} = \{I_{John,\Psi}, I_{John,k}\},$$

$$\neg I^S(I_{John,\Psi K}, e_{sdb})$$

And in relation with e_{ins}

$$\neg I^S(I_{John,k}, e_{ins}),$$

$$\neg I^S(I_{John,\Psi}, e_{ins}),$$

$$I^S(I_{John,O\Psi}, e_{ins}),$$

$$\mathrm{I}^{\mathrm{S}}\left(I_{John,O\mathrm{k}}, \mathrm{e}_{\mathrm{ins}}\right)$$

The region information as a single information is not sensitive with *sharing* with $\mathrm{e}_{\mathrm{ins}}$.

$$\neg\mathrm{I}^{\mathrm{S}}\left(\mathrm{I}_{\mathrm{John},\Upsilon}, \mathrm{e}_{\mathrm{ins}}\right)$$

In result of a scientific research, the correlation of K and Ψ also is not sensitive:

$$\neg\mathrm{I}^{\mathrm{S}}\left(I_{John,\Psi\mathrm{K}}, \mathrm{e}_{\mathrm{ins}}\right)$$

Applying differential privacy can limit the *disclosure* of $I_{John,O\Psi}$ and $I_{John,O\mathrm{k}}$ to other entities including $\mathrm{e}_{\mathrm{ins}}$.

However, another processing operation $\mathrm{o}_{\mathrm{sdb,reg}}$ in $\mathrm{e}_{\mathrm{sdb}}$ is applied on information to evaluate the correlation of regions and the disease. The outcome may conclude that 99 percent of people having Ψ are coming from Υ. Although Υ is not considered being sensitive, the combination of the disease and the region becomes sensitive.

$$I_{John,\Upsilon\Psi} = \left\{\mathrm{I}_{\mathrm{John},\Upsilon}, \mathrm{I}_{\mathrm{John},\Psi}\right\}$$

$$\mathrm{I}^{\mathrm{S}}\left(I_{John,\Upsilon\Psi}, \mathrm{e}_{\mathrm{ins}}\right)$$

Regardless of John's interests in participating with the statistical database towards achieving a treatment for Ψ, he does not agree on participating in $\mathrm{o}_{\mathrm{sdb,reg}}$ (though it may have a low L1-sensitivity measure). The mechanism in differential privacy does not provide the autonomy for entities to evaluate if the existing operations are authorized to be applied on their information.

Lack of sensitivity concept at the information level makes differential privacy fail in various scenarios in the setting of statistical databases. This model has motivated several privacy models and privacy mechanism designs such as [63], [32], [64]. Similar to differential privacy, any model that can be reduced at information management can be explained and abstracted by the proposed framework in which supports diverse settings of privacy among entities.

3.3 CDS: Adequate privacy model

The analysis within our research indicates that among the existing privacy models, attending to settings that can be adequate for CDS environments is lacking. The privacy model in CDS has to be captured at the computation and therefore, it requires formal modeling of privacy. We claim that the proposed privacy model that we presented in chapter 5 is associable and applicable in CDS environment as computation platforms.

The proposed formal privacy model is in the context of information management where entities are modeled as set of information and operations. Information management is categorized as information collection, processing and dissemination.

CDS is a class of systems that is positioned as a computation platform in which computation happens at the interactions of entities. Solutions in CDS are achieved by participation of entities in a distributed decentralized fashion. They require resolving the interdependency problem through coordinating their activities for which they adopt interaction mechanisms.

In incomplete knowledge world, entities' knowledge about the world is incomplete for which entities update their knowledge about the world and solve their problems through message-based interactions.

DEFINITION 1: A computation system including entities E provides a solution (S) to a problem (P) by applying computation processes (Cp).

$$C: E \times P \times Cp \rightarrow S$$

DEFINITION 2: C is Information Management computation system $(i - mng)$ when problem and solution are modeled as information and computation as operation.

Operations in information management can be classified as collection, processing and dissemination that can be executed by entities E. If the problem is modeled as information I:

$$\forall\, i, \in (i, P) \wedge \exists\, s\,, \in (s, S)\ \wedge \exists\, o \in \left(o, O_j\right) \wedge \ni\ (e_j, E)| = (\ o(i), s) \rightarrow\ \ C: E \times I \times O \rightarrow S$$

DEFINITION 3: S is acceptable solution $(s - accept(s))$ is it resolves the problem and does not result in privacy concern.

DEFINITION 4: Privacy Model in the context of sensitive information $(P - Model)$ is

$$P: \{e_i, e_j\} \times I_i \times O_j \rightarrow \bigcup_{k=1}^{k \leq M} \neg I^s(I_{i,k}, e_j)$$

THEOREM 1: Let P be a $(P - Model)$. For any $(i - mng)$, P is essential to have $(s - accept)$.

$(P - Model)$: $\{e_i, e_j\} \times I_i \times O_j \rightarrow \bigcup_{k=1}^{k \leq M} \neg I^s(I_{i,k}, e_j)$

$(i - mng)$: $C: E \times I \times O \rightarrow S$

$$\forall\, i, j, k \mid Q \equiv \bigcup_{k=1}^{k \leq M} \neg I^s\left(I_{i,k}, e_j\right)$$

$$if\ \exists\ s, \in (s, S) |\ \notin (s, Q) \rightarrow\ \neg(s - accept(s))$$

Therefore, the acceptable solution in $(i - mng)$ has to include the $(P - Model)$.

THEOREM 2: Any incomplete knowledge CDS computation is an $(i - mng)$

The computation in incomplete knowledge CDS happens in interactions therefore:

$$C: E \times In \rightarrow S$$

Because in incomplete knowledge CDS, knowledge in modeled as information, interaction is modeled as information collection, processing and dissemination which can be abstracted as Operation and information. Hence:

$$In\ \equiv < I, O >$$

$$C: E \times In \rightarrow S$$

$$C: E \times I, O \rightarrow S$$

Therefore, computation in incomplete knowledge CDS can be modeled as information management computation, which based on THEOREM 1 affirms the proposed privacy model is applicable and required to achieve acceptable solutions.

3.4 Summary

Privacy can adequately be addressed in the context of information management where the information is collected, processed and disseminated. Entities tend to protect their sensitive information by not *sharing* it with other entities. Sensitivity is a relative concept that may change from a receiver to another. For any information, there exists an exposure boundary that includes entities that the flow of information within the boundary is not a concern. However, transferring information to outside of the boundary through operations makes the information sensitive and cause privacy concern. The information can be classified as explicit and implicit. The latter one is the conjunction of explicit information and operations. The privacy concern is related to the transformation of the explicit information to sensitive information by applying operations. Such operations become no-authorized operations. If there is an agreement between the entities on not executing the non authorized operations, and still the operation is executed, privacy has been violated. In addition, if an operation is applied that prevents or neutralizes the application of information, it is referred as privacy protection. Many privacy models can be modeled through the proposed privacy model and it is illustrated that there are sensitive information that may not be protected within those models. Furthermore, it is formally argued that the privacy model at information management context can adequately address privacy in CDS environments.

Chapter 4

4 Privacy Protection Management Framework

The computation in CDS happens at interaction level where entities exchange information at which information management becomes an adequate context to model privacy for CDS. Additionally, the message-based interactions in CDS can be modeled with information management into information collection, processing and dissemination. Modeling privacy in information management capacitates application of the model in interactions through which privacy becomes part of the computation. The interactions are steered by interaction protocols that are abstracted as set of messages and sequences. By incorporating the privacy model at the interaction, it creates a privacy protection management framework that can expand on interaction protocol messages and sequences that are supported by privacy protection mechanisms.

4.1 Our Contribution

Interaction-Based Privacy Protection Management Framework: In this work we have proposed an interaction-based privacy protection framework that includes the formal analysis of capturing the privacy requirements in interaction of entities as well as applying adequate privacy protection mechanisms. As the computation in CDS happens at interaction, the proposed expands the interaction protocol with privacy protection mechanisms. The proposed framework is an architectural based solution for privacy and is defined at the interaction level. This enables the framework to be adopted by various applications and computational solutions.

Analytical Tool for elaborating privacy as a state in a computational system: The proposed privacy model and the privacy protection management framework provide the necessary tools to evaluate the state of privacy in different systems by processing on the interaction protocol that is applied. We have used this framework to identify the privacy concerns at Contract Net protocol that is discussed in more details in Chapter 6.

Capturing privacy as a computation concept. By applying the privacy protection management framework at computation level, the privacy protection management becomes an architectural element of a computation entity as well as a new parameter essential for computation solution functions. The details of implementing this framework at computation entity are provided in Chapter 5.

Privacy-Based Interaction Protocol. The proposed privacy protection management framework expands the given interaction protocol with adequate privacy protection mechanisms which can provide sufficient privacy protection for entities in CDS.

Quantifiable Privacy Protection Level for the privacy based interaction protocol. The proposed privacy protection framework conducts analysis on measuring the privacy protection level of protection mechanisms that leads to measuring the protection level of the provided privacy based interaction protocol.

4.2 Privacy Protection in Incomplete Knowledge in CDS

In order to manage privacy protection, privacy protection mechanisms require knowing the operations of entities and being aware of what operations are authorized. In various instances of CDS environments, knowledge of entities is incomplete. This implies uncertainty about operations of entities. Capturing uncertainty provides levels of knowledge about the operations. This affirms the exercise of quasi protection mechanisms in varied precedent of CDS environments.

Quasi protection mechanisms convey levels of uncertainty about the extent of unauthorized operations that the mechanism can prevent from execution. For instance anonymization techniques can provide privacy protection with a degree of probability [23], [65]. Others such as rule based mechanisms for protecting privacy are capable of supporting a limited number of non-authorized operations [9], [42]. The uncertainty level in these cases has been captured as Privacy Protection Level (PPL). In the other word, PPL is a probabilistic base model to describe the effectiveness of a mechanism to prevent or neutralize unauthorized operations from producing sensitive information. This measure can be associated to computational concepts. The execution of the mechanism μ in

relation to protecting privacy is the space S that the mechanism can prevent the execution of non-authorized operation:

$$\bar{\bar{\mu}} \equiv PP\left(e_j,\ S, \hat{O}_j^i(S)\right)$$

By applying the mechanism over the space of entities' information set, there is uncertainty level associated to the application of the protection mechanism which implies the conditional probability protecting privacy by executing μ given the space of I_i. In another word, we measure the probability of μ protecting privacy when it is applied on I_i.

$$PPL(e_j, I_i, \mu_t) = P(\ \bar{\bar{\mu}} \mid I_i) \quad 9.$$

This can be measured either statistically or characteristically. For instance, in a simplified view, in a complete knowledge world where entities have the knowledge over all communicated information, in discrete set of operations and an algebraic form, evaluating PPL depends on non-authorized operations that is prevented from application by applying the mechanism z to all of non-authorized operations n; PPL $= \frac{z}{\mathrm{n}}$.

PPL is a measure that predicts privacy protection in an interaction among two entities. Evidently, if the mechanism can provide outputs where$(z = n)$, perfect privacy protection is achieved.

Depending on the context and architecture of the environment, PPL might be evaluated differently using the same approach. As an example, in this section, we evaluate the PPL of differential privacy [23]. A randomized function K is $\epsilon-$ differentially private if for all datasets $\mathrm{D}_2$and D_1differing on at most one element and all $\mathrm{S} \subseteq \mathrm{Range(K)}$, $\Pr\left[\mathrm{K}(\mathrm{D}_1) \in \mathrm{S}\right] \leq \exp(\epsilon) \times \Pr[\mathrm{K}(\mathrm{D}_2) \in \mathrm{S}]$[23].

To achieve differential privacy, a mechanism is required that can implement differential privacy [66], [63]. The probability of a mechanism implementing differential privacy is $1 - 2\epsilon$ [63].

Considering n as number of non-authorized operations [queries] in info collector, implementing $\epsilon - differential\ privacy$ in z number of non-authorized operations has $1 - 2\epsilon$ probability in each of them. Therefore, it creates a binomial distribution in which the expected value of z: $E(z) = n(1 - 2\epsilon)$. This leads to $PPL = 1 - 2\epsilon$.

4.3 Privacy protection mechanism

As described in previous sections, privacy protection mechanisms apply on operations that prevent or neutralize non-authorized operations. Privacy protection mechanisms in both forms of perfect or quasi can effect in two dimensions. Either they can work at the operation level to identify the non-authorized operations such as rule based authorization engines or they can apply at the information level to neutralize the execution of non-authorized operations such as distorting the information as results of operations.

Privacy protection mechanisms are operations that are applied on information and provide the necessary information for privacy protection. This indicates that the structure of privacy protection mechanism is the set of operations it applies O^{μ} and set of information generated by the operations I^{μ}.

$$\mu =< O^{\mu}, I^{\mu} > \text{ 10.}$$

$$O^{\mu} = \{o^{m,1}, \dots, o^{m,d}, \dots, o^{m,D}\}, 1 \leq d \leq D$$

Privacy protection mechanism also can be categorized as preventive and punishing. When the mechanism operations is applied before *sharing* information, it is preventive and when it is practiced after non-authorized operations are executed, they become punishing mechanisms.

Preventive mechanisms at the information level refer to protection mechanisms that are running in a sequence to provide sufficient information for the requested service or task in addition to not *disclosing* the sensitive information. In this context, there exist at least two entities that one of them owns the information e_i . The other entity e_j is collecting information i to perform a service or a task. Therefore, the information i has to be *shared* with the collecting entity.

$$\mathrm{S}\left(\mathrm{I}_{i,\mathrm{k}}, \mathrm{e}_j\right)$$

Applying the protection mechanism at preventive at information level would be as the following:

$$\forall\, o^{m,d} \in\ O^{\mu} |\ Subset(\bar{\bar{o}}^{m,d}\left(\mathrm{I}_{i,\mathrm{k}}\right),\ I^{\mu})$$

$$\bar{\bar{o}}^{m,1}\left(\mathrm{I}_{i,\mathrm{k}}\right) = \mathrm{I}_{i,\mathrm{k}'} \wedge (\nexists o^{m,d'} \in O^{\mu} | o^{m,1} \equiv o^{m,d'})$$

$$\bar{\bar{o}}^{m,D}\left(\bar{\bar{o}}^{m,D-1}\left(\bar{\bar{o}}^{m,D-2}\left(\mathrm{I}_{i,\mathrm{k}}\right)\left(\bar{\bar{o}}^{m,D-3}\left(\mathrm{I}_{i,\mathrm{k}}\right)\left(\ldots\left(\bar{\bar{o}}^{m,1}\left(\mathrm{I}_{i,\mathrm{k}}\right)\right)\right)\right)\right)\right) = I^{\mu} \quad 11.$$

$$\wedge\ (\nexists o^{m,d"} \in O^{\mu} | o^{m,D} \equiv o^{m,d})$$

Every operation in the mechanism provides the information for another operation in the mechanism. These operations also will be executed in a specific pattern of sequence. The examples of these mechanisms are anonymization techniques [23], [43], [13], [12] or encryption methodologies [56], [59], [60] applied for privacy protection.

Similarly, preventive privacy protection mechanism at the operation level includes operations that are performed in a specified order. Executing these operations does not allow the non-authorized operations retrieve any result.

$$\forall\, t\ |\widehat{O}_{j,t}^{i,k}(\mathrm{I}_{i,\mathrm{k}}) \equiv \mathrm{I}_{i,\mathrm{k}'}\ \wedge \mathrm{I}_{i,\mathrm{k}'} \notin I^{\mu}$$

Operations in this type of protection mechanisms require being aware what operation is going to be applied on information. They either are not authorized and therefore, do not get results or the result will be provided for them.

$$\bar{\bar{o}}^{m,D}\left(\left\{o_{j,w},\ \bar{\bar{o}}^{m,D-1}(\{o_{j,w},\ \bar{\bar{o}}^{m,D-2}(\{o_{j,w}, \ldots, \bar{\bar{o}}^{m,1}(o_{j,w}, \mathrm{I}_{i,\mathrm{k}})\})\})\right\}\right) = \begin{cases} \emptyset & if \in (o_{j,w}, \widehat{O}_j^i) \\ \mathrm{I}_{i,\mathrm{k}"} & if \notin (o_{j,w}, \widehat{O}_j^i) \end{cases}$$

12.

The punishing approach in privacy protection mechanism is applied in situations where the prevention of *sharing* information is not possible. However, some operations can provide assurances to the owner of information. Whenever a collecting entity violates their privacy requirements, the owner of the information can exercise some degree of authority in executing the punishing operations accordingly. The example of this approach is the terms and conditions that are accepted by both entities. If any operation outside of the agreement is executed, there will be legal consequences for the non-compliant entity. The generated information in this mechanism is *shared* with the entity that has executed the non-authorized operations.

$$\forall\, t, \widehat{o}_{j,t}^{i,k} \mid \overline{\overline{\mathrm{o}^{\mathrm{m,D}}}}\left(\left\{\widehat{o}_{j,t}^{i,k},\ \overline{\overline{\mathrm{o}^{\mathrm{m,D-1}}}}(\{\widehat{o}_{j,t}^{i,k},\ \overline{\overline{\mathrm{o}^{\mathrm{m,D-2}}}}(\widehat{o}_{j,t}^{i,k}, \dots, \overline{\overline{\mathrm{o}^{\mathrm{m,1}}}}(\widehat{o}_{j,t}^{i,k})\})\})\right\}\right) \equiv I^{\mu} \quad 13.$$

The classification of protection mechanism are depicted in Figure 2.

When preventive mechanism cannot be applied, the punishing mechanism will be more adequate. For instance, when a service provider interacts with a consumer in different time periods, the information that are aggregated in this period can be used to transform sensitive implicit information to explicit using an auxiliary information. To avoid this, punishing mechanism will be more effective. Naturally, punishing mechanisms support agreements among two entities in which enforce the execution of consecutive actions towards the faulty entity.

Protection mechanism can be applied at information and operation levels. Typically protection mechanisms at the information level limit the access of entities to the information that is *shared*. As an example anonymization and encryption distort the information for which is sufficient for resolving the requested task and does not *disclose* the sensitive information. This might be inadequate in relation with applications that require receiving the non-distorted complete information. To deal with this the protection mechanisms at the operation level are more advantageous.

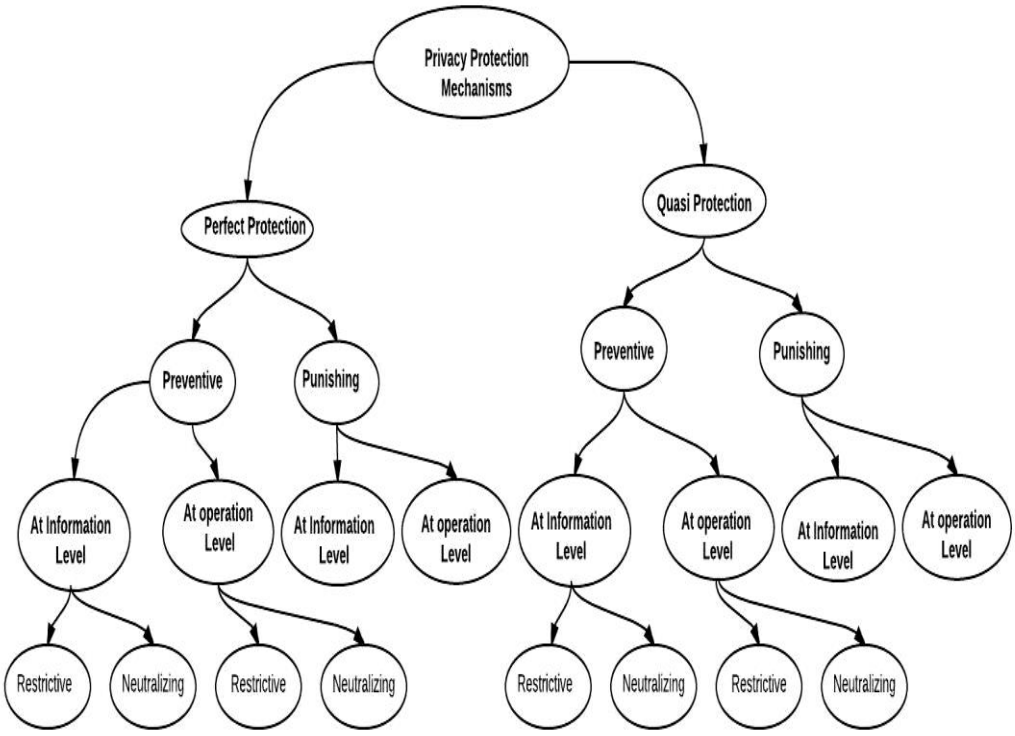

Figure 2. Classification of protection mechanisms

4.4 Privacy Protection Management Framework

Typically, entities have various expectations and preferences over privacy that by itself changes in different contexts and in different times. This is the reason that makes rule-based protection mechanism not convenient for CDS environments. These approaches are well established where the social and technological developments are in a stable state [10, 40] . In contrary of rule based approaches, there are standard based approaches that are commonly applied at the architecture level. One of the instances of standard based approaches is PII 2.0. In this model, the risk of interaction of entities is a measure to decide proceeding interactions. If the risk of interaction is not acceptable by the entity, the refuse or look search for alternatives otherwise, they take the risk and *share* the required information [40]. This motivated us to work on a framework that can evaluate

the risk of interaction and possible privacy protections to enable entities making decisions that can protect their privacy in addition to resolving the interdependency problem.

4.5 Privacy Protection at the interaction level

In CDS, privacy can be reduced to operations and information which enables it being part of information management. Information management also can be categorized as information collection, information processing and information dissemination. To employ information management, it can be carried at the interaction level where the information is collected, disseminated or processed. Providing privacy protection at the interaction level is an architectural approach that can benefit various applications at entities that are using interaction protocols to resolve their interdependency problem.

Entities might need to exchange information while they are interacting. This is the initial point where the information is *shared*. However, it also emphasizes on the focus of this work which is on message based interactions. Providing the privacy protection mechanism at the interaction protocol enables applications on entities to delegate the privacy resolution procedure to the interaction protocol and the solution space of those applications will be limited to entities that can protect entities privacy.

4.5.1 Privacy-based interaction protocol

Depending on the type of interdependency, interaction can be modeled as the following:

$$Interaction = \ < \delta,\ e_i,\ e_j, IP >$$

δ is the type of interdependency [19], e_i is the entity that requires capabilities from an entity such as e_j. IP is the interaction protocol acquired by entities to coordinate their activities. Message based interaction protocols can be modeled as a set of messages and the pattern of sequences that includes messages that is exchanged among entities.

$$IP =< M, S(M) >$$

In the previous tuple, M is the set of messages and $\mathrm{S}(M)$ is denoting the sequences that are generated by the protocol.

Sequences in the interaction protocol refer to the pattern of the exchanged messages. In fact, a given sequence indicates where information is collected and disseminated. As described in the proposed privacy model, collecting and disseminating information can be reduced at the operation level. Similarly, the existing sequences of an interaction protocol also can be modeled by sequence of operations. Therefore, the structure of interaction protocols can be reduced to operations and be modeled as:

$$IP = [o^{IP,1}, \dots, o^{IP,q}, \dots, o^{IP,Q}] \quad 14.$$

To protect privacy at the interaction level, privacy protection mechanism should be incorporated with the operations of interaction protocol. As discussed, privacy protection mechanisms have set of operations that are executing in a specific order:

$$O^{\mu} = [o^{m,1}, \dots, o^{m,d}, \dots, o^{m,D}] \quad 15.$$

Each of the operations of the privacy protection mechanism might belong to an entity that exists in the environment. Our assumption is that the privacy protection mechanism involves entities that match with the architecture of the interaction protocol.

The privacy protection management framework requires transforming the interaction protocol to a protocol that is integrated with privacy protection mechanisms and delivers the solution it is designed for. One of the objectives of the proposed framework is to provide a solution space that meets the privacy requirements. To achieve this, the framework merges the operations of the privacy protection mechanism with the operations of the interaction protocol in a totally ordered fashion. It can happen in three forms; either the protection mechanism operations is concatenated to the list of interaction protocol operations as prefixes

$$[o^{m,1}, \ \dots, o^{m,D}, o^{IP,1}, \dots, o^{IP,Q}]$$

Such as applying anonymization and encryption operation before *sharing* the information with a service provider; or appended to them such as the operations that happens by re-enforcements:

$$[o^{IP,1}, \dots, o^{IP,Q}, o^{m,1}, \dots, o^{m,D}]$$

It also can be merged in with other interaction protocol operations in a way that the order of interaction protocol operations does not change and the order of mechanism operations does not change yet they intervene. The location of intervention will be specified by evaluating the message to verify if it belongs to sensitive information.

$$[o^{IP,1}, \dots, o^{m,1}, \dots, o^{IP,q}, \dots, o^{m,d}, \dots, o^{IP,q'}, \dots, o^{m,d'}, \dots, o^{IP,Q}]$$

By capturing the exposure boundary, it is possible to identify the sensitive information. If information belongs to sensitive information, there is a protection mechanism that can prevent the execution of non-authorized operations on them. Therefore, any operation in interaction protocol that *discloses* the sensitive information will be substituted with sequences of operations that include the protection mechanism.

Given the operations in IP and operations in protection mechanism, every operation in protection mechanism has been targeted for protecting a sensitive information.

$$\forall o^{m,d}, \exists I_{i,k} | \mathrm{I}^{\mathrm{S}}(\mathrm{I}_{i,k}, \mathrm{e}_j) \rightarrow \exists o^{IP,q}, I_{i,k'} | o^{IP,q}(I_{i,k'}, I^{aux}) \equiv D(I_{i,k'}, e_j) \wedge$$

$$\bar{\bar{o}}^{m,d} \rightarrow \neq (\bar{\bar{o}}^{IP,q}(I_{i,k'}, I^{aux}), \mathrm{I}_{i,k})$$

Therefore, any operations in interaction protocol that *discloses* the sensitive information will be replaced by the sequence of the mechanism operation and interaction operation. It can happen at two levels either at preventive level that can be prefixing:

$$\overrightarrow{Seq}(o^{IP,q}, o^{m,d}) = [o^{IP,q}, o^{m,d}]$$

or at the punishing level that can be appending:

$$\overrightarrow{Seq}(o^{IP,q}, o^{m,d}) = [o^{m,d}, o^{IP,q}]$$

Merging of the operations of privacy protection mechanism with operations of interaction protocol requires extending the message types and sequences of the protocol as explained

above. This becomes the extended interaction protocol that integrates the privacy protection mechanism at the interaction level.

Let's R_i be the exposure boundary including information about information in e_i . Based on the information that is *shared* through the interaction protocol and the R_i, there is a protection mechanism that can prevent execution on non-authorized operations.

The proposed framework using the provided information at the risk evaluation, PPL evaluation and the interaction protocol reduces the space of possible solutions to ones that can provide the privacy protection expected. It can be modeled as

$$PrivacyProtectionMngFRMK = < interaction,\ RiskEvalution,\ PPLEvaluation,\ PBIP >$$

Figure 3 depicts the operational view of the privacy protection management framework. By applying the risk evaluation model, it is possible to identify the *sensitive* information that might be *shared* among entities of the environment while interacting. The messages and sequences of messages among entities construct the interaction protocol of that

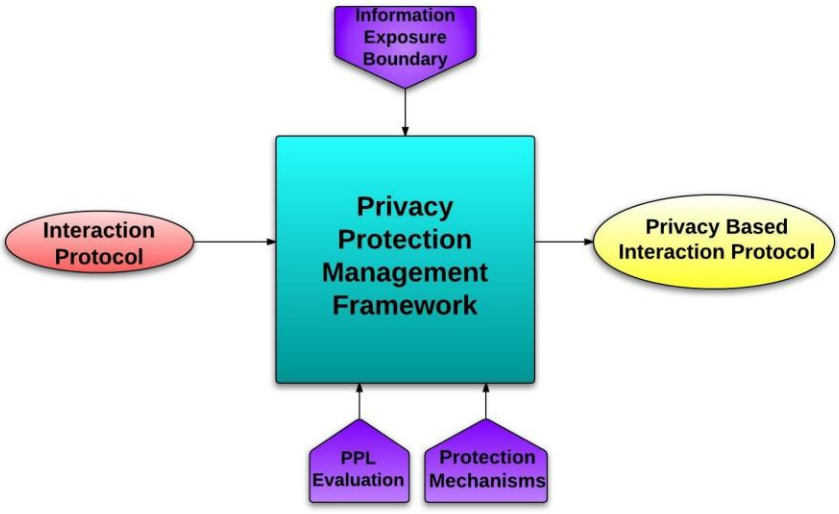

Figure 3. Operational view of privacy protection management framework

environment. Providing the framework with the exposure boundary, the interaction protocol, PPL evaluation and the type of privacy protection mechanisms the framework can provide messages and sequences that represent the privacy based interaction protocol. Entities that adhere to this interaction protocol seamlessly interact with other entities and the interaction protocol applies the privacy protection operations to protect privacy independent from the application. This allows the privacy protection in CDS be incorporated at the architectural level and part of the computation platform.

The operations in the privacy protection mechanism may require new type of messages in the message set of the protocol in addition to the extension on the sequence of interaction protocol. Through accommodating privacy protection mechanism at the interaction protocol level, the interaction is limited to entities that privacy can be protected with an acceptable PPL in their interaction. The sequence of the operations in interaction protocol is not changed in the privacy based interaction protocol but the operations of the privacy protection mechanisms are applied. This can prevent or neutralize execution of non-authorized operations and transforming the sensitive implicit information to explicit. Each of the applied mechanisms has a PPL value. Several mechanisms can be integrated with an interaction protocol to form a privacy based interaction protocol (PB_IP). By putting an assumption on independency of the protection mechanisms, the PPL of the protocol becomes the multiplication of PPL of all applied mechanisms.

$$PPL\left(R^{*}, \bigcup_{i=1}^{N} I_i \mid \in (e_i, R^{*}), PB\_IP\right) = \coprod_{\forall\, \mu, j \mid \in(\mu, PB\_IP), \in(e_j, W)} PPL(e_j, \mathrm{M(PB\_IP)}, \mu) \qquad (16)$$

The privacy protection management framework identifies the sensitive information by capturing SO information and their exposure boundary. This enables the framework to identify what privacy mechanisms at the interaction protocol level is required. However, the privacy based protocol extends the given interaction protocol without altering the mechanism that the interaction protocol employ to resolve the interdependency problem. Also, the incorporated privacy protection mechanisms does not change the architecture of

interdependency resolution mechanisms but extends the protocol with adequate messages and sequences to provide privacy protection with certain PPL level.

4.5.2 Privacy Protection Management Framework

In this work, we developed a privacy protection management framework that is applied at the interaction level where the interaction protocol is expanded with additional operations to support privacy protection. There are three theorems within the proposed framework that we demonstrate the related formal proof in this section.

4.5.3 Privacy can sufficiently be protected at the interaction level

The proposed framework provides the protection mechanisms at the interaction level. It extends the interaction protocol with essential messages and sequences to protect the sensitive information that is shared or disclosed in the original interaction protocol.

Theorem 3: For any incomplete knowledge CDS where entities adopt message-based interaction, P-Model can be sufficiently addressed at the interaction level.

To provide the supporting materials for the above theorem, it is essential that we prove the following points:

- All the information that is *shared* or disclosed to other entities are decided at the interaction level
- Any class of privacy protection mechanism happens at the interaction level.

The computation entity in CDS has autonomy on coordinating activities with others. The interaction layer manages the necessary processes to identify the adequate messages to communicate to resolve the interdependency problem. The communication layer is responsible for exchanging messages. However, it does not have the decision-making authority on the messages to be sent and it is not aware of the intent that initiates the exchange of messages.

Proof:

Lemma 1: Let $e_i \equiv < K_i, PS_i, In_i, Com_i >$ *be the computation entity. For any information* $I_{i,r}$ *that is going to be shared with* e_j, $S(I_{i,r}, e_j)$ *is decided in* In_i

If PS_i realizes that to achieve a goal, there is interdependency problem, In_i finds a coordination solution $CS_{i,j}$ with an entity such as e_j .

If $I_{i,r}$ is *shared* with e_j,

$\exists\, I_{i,r}, \in (I_{i,r}, I_i^k) | S(I_{i,r}, e_j)$

Then there are two possibilities:

1. It is discovered at PS_i that $I_{i,r}$ is required to perform the $CS_{i,j}$ therefore

   $CS_{i,j} \rightarrow S(I_{i,r}, e_j)$

2. It is discovered at In_i that $I_{i,r}$ has to be *shared* with e_j

   $In_i \rightarrow S(I_{i,r}, e_j)$

In both cases, the *shared* information is processed and decided by the interaction layer.

Lemma 2: Let $I_{i,r}$ be the information that is disclosed. For any $I_{i,r}$ there is explicit information that is *shared*

$\exists\, I_{i,r}, e_j\, , \in (I_{i,r}, I_i^k) |\, D(I_{i,r}, e_j)$

When information is implicitly disclosed:

$D(I_{i,r}, e_j) \rightarrow \exists\, I_{i,r'}, o_{j,w} |\, o_{j,w}(I_{i,r'}, I^{aux})$

Assuming $I_{i,r'}$ is not *shared* through interaction. Then there are two possibilities:

1. Fact A: $I_{i,r'}$ is an auxiliary information disseminated by a third party e_t then:

   Using lemma 1:

If $I_{i,r'}$ is *shared* to e_t, then it has been decided at interaction

2. $I_{i,r'}$ is not *shared* with any entity, therefore:

    a. Either $D(I_{i,r'}, e_t)$ so that Fact A occurs

    b. Or it has not been *shared* by interaction. This contradicts Lemma 1.

This proves that any information that is *shared* or *disclosed* has initiated *sharing* point at the interaction.

In equation 8, Privacy protection in privacy model is defined as :

$$\mathrm{PP}\left(\mathrm{e_j},\ \left(\mathrm{PS(I_i)}\right), \widehat{\mathrm{O}}_\mathrm{j}\right) \equiv \forall\, \mathrm{t}\,, \mathrm{w}|\ \subset (\mathrm{t}, \mathrm{PS(I_i)})\ \wedge\ \widetilde{\bar{\bar{\delta}}^{\mathrm{t}}_{\mathrm{j,w}}(\mathrm{t})}$$

To achieve $\widetilde{\bar{\bar{\delta}}^{\mathrm{t}}_{\mathrm{j,w}}(\mathrm{t})}$, the privacy protection mechanisms are applied. The privacy protection mechanisms can be classified at information or operation level.

Lemma 3: If a preventive protection mechanism at information exists, it happens at the interaction.

Let μ be a preventive mechanism at information level for protecting $I^s(I_{i,r}, e_j)$ in which enables $\widetilde{\bar{\bar{\delta}}^{\mathrm{t}}_{\mathrm{j,w}}(\mathrm{t})}$.

$$\bar{\bar{\mu}} \rightarrow\ \mathrm{PP}\left(\mathrm{e_j},\ \{\mathrm{I_{i,r}}\}, \widetilde{\bar{\bar{\delta}}^{\mathrm{t}}_{\mathrm{j,w}}(\mathrm{t})}\right)$$

In Equation 10,

$\mu \equiv < I^{\mu}, O^{\mu} >$

$O^{\mu} = \{o^{m,1}, \dots, o^{m,d}, \dots, o^{m,D}\}, 1 \leq d \leq D$

Based on the execution of preventive protection mechanisms at information level in equation 11:

$$\bar{\bar{o}}^{m,D}\left(\bar{\bar{o}}^{m,D-1}\left(\bar{\bar{o}}^{m,D-2}(\mathrm{I}_{i,\mathrm{k}})\left(\bar{\bar{o}}^{m,D-3}(\mathrm{I}_{i,\mathrm{k}})\left(\ldots\left(\bar{\bar{o}}^{m,1}(\mathrm{I}_{i,\mathrm{k}})\right)\right)\right)\right)\right) = I^{\mu}$$

This results in sharing information that is manipulated by the operations in protection mechanisms.

$\bar{\bar{\mu}} \rightarrow S(I_{i,r'}, e_j)$

Based on Lemma 1, $I_{i,r'}$ has to go through interactions. Therefore, the preventive mechanisms at the information level can happen at the interaction level.

Lemma 4: If a preventive mechanism at operation level exists, it happens at interaction level

Let $I_{i,r}$ be the sensitive information that can implicitly be *disclosed* to e_j through $\hat{o}^t_{j,w}$ when $I_{i,r'}$ is *shared*.

$$\exists\, I_{i,r}, I_{i,r'}, e_j, \hat{o}^t_{j,w} |\; I^s\left(I_{i,r}, e_j, \hat{o}^t_{j,w}\right) \wedge \hat{o}^t_{j,w}(I_{i,r'}, I^{aux}) \equiv I_{i,r} \wedge S(I_{i,r'}, e_j)$$

Let μ be the protection mechanism at the operation level that can protect $I_{i,r}$.

$\mu \equiv < I^{\mu}, O^{\mu} >$

$O^{\mu} = \{o^{m,1}, \ldots, o^{m,d}, \ldots, o^{m,D}\}, 1 \leq d \leq D$

Based on the execution of the protection mechanisms at the operation:

$$\bar{\bar{o}}^{m,D}\left(\left\{o_{j,w},\ \bar{\bar{o}}^{m,D-1}(\{o_{j,w},\ \bar{\bar{o}}^{m,D-2}(\{o_{j,w}, \ldots, \bar{\bar{o}}^{m,1}(o_{j,w}, \mathrm{I}_{i,\mathrm{r}})\})\})\right\}\right) = \begin{cases} \emptyset & if \in (o_{j,w}, \hat{O}^i_j) \\ \mathrm{I}_{i,\mathrm{r}''} & if \notin (o_{j,w}, \hat{O}^i_j) \end{cases}$$

which results in *sharing* $\mathrm{I}_{i,\mathrm{r}''}$ or $\emptyset$. Therefore, based on Lemma 1, it happens at the interaction level.

Lemma 5: if there is punishing privacy protection mechanisms, it happens at the interaction level.

Let $\mu \equiv < I^{\mu}, O^{\mu} >$ be the punishing protection mechanism that protects $I^{s}(I_{i,r}, e_j)$. Based on the execution of punishing mechanisms in equation 13:

$$\forall\, t, \widehat{O}_{j,t}^{i,k} \mid \overline{\overline{\mathrm{o}^{\mathrm{m,D}}}}\left(\left\{\widehat{O}_{j,t}^{i,k}, \overline{\overline{\mathrm{o}^{\mathrm{m,D-1}}}}(\{\widehat{O}_{j,t}^{i,k}, \overline{\overline{\mathrm{o}^{\mathrm{m,D-2}}}}(\widehat{O}_{j,t}^{i,k}, \dots, \overline{\overline{\mathrm{o}^{\mathrm{m,1}}}}(\widehat{O}_{j,t}^{i,k})\})\})\right\}\right) \equiv I^{\mu}$$

The generated information in this mechanism is *shared* with the entity that has executed the non-authorized operations.

$\bar{\bar{\mu}} \rightarrow S(I^{\mu}, e_j)$

This indicates that the punishing mechanisms happen at the interaction level.

Given Lemma 1, Lemma 2, Lemma 3, Lemma 4 and Lemma 5, it is proven that any protection mechanisms will be applied at the interaction level. Therefore, it is sufficient to capture the privacy protection at the interaction level.

4.5.4 Quantifiable Protection in Privacy-Based Interactions

The privacy protection mechanisms can be classified as preventive at information, preventive at operation and punishing mechanisms. In this section, we argue that for each of the privacy protection mechanisms at information management, there are degrees of probabilities for privacy protection when they are applied at the context of information management.

Privacy protection mechanism at the information level attempts to manipulate the given information. Such a characteristic makes the mechanism meaningful enough for the service operations, and at the same time, it is a desirable choice as it does not disclose the sensitive information. Typically, this happens by distorting the information by adding noise or altering through particular formats. The major classes of mechanisms in this category are anonymization and cryptographic methods. Anonymization methods are associable with a degree of confidence factor that reflects the effectiveness of the

anonymization function and the degree of re-identifying. For instance in ε −differential privacy, higher level of ε has lower confidence factors in de-identifying and lower values of are more effective in anonymizing information of the dataset [62]. In approaches such as k-anonymity, l-diversity and t-closeness the parameter k, l and t reflect on the capability of the anonymization function on the de-identifying information. Also, cryptographic methods are mainly related to the difficulty of breaking the key that is used for altering the information that is encrypted. However, based on the length of the key that is used for encryption, the exhaustive methods in brute force fashion theoretically can break the code. This will position the probability of cryptography systems is a function of the length of the key.

The preventive privacy protection mechanisms at operation level can include contractual cryptographic mechanisms, altering the non-authorized information and rule-based mechanisms. There are efforts in formalizing the rule sets that are applied in privacy mechanisms[22]. This allows analyzing information that will be protected by the privacy protection rules. Furthermore, in a statistical analysis for a given mechanism such as policy-based mechanisms, the probability distribution of the mechanism, it would be possible to measure the probability of privacy protection of the mechanism.

The punishing privacy protection mechanisms are tailored to applying operations that negatively impact the utility of the entity that executed the unauthorized operations. For instance, the reputation systems and legal consequence of agreement violation affects regressively on the utility of the entity if it exceeds the agreement. The study of the impact of punishing mechanisms on the decision-making process of an entity to execute unauthorized operations are captured in utility theories and economic mechanisms that are part of the future works of the current research.

Through this, the privacy protection level of the interaction protocol that adopts the privacy mechanisms in the context of information management is quantifiable.

For example, incorporating anonymization mechanism such as differential privacy, encrypted bid submission, early registration and terms and condition types of agreement in the PB_CNP will provide the following PPL:

Assuming the information is segregated to the lowers possible granularity, the protection mechanisms become independent from each other. Therefore:

- Anonymization Differential Privacy: the probability of creating a differential private function is $(1 - 2\varepsilon)$

- Privately-Communicate-bid: the work in [57]has applied the public private key cryptographic mechanism to communicate the bid value. Assuming the length of the key is , the probability of breaking the code would be $\frac{1}{2^n}$. In practice, breaking the cryptographic mechanisms that have high length key, it is close to impossible. However, collusion of the participant entities may circumvent the encryption mechanism. Assuming the probability of collusion among the manager and the bidders in p'. Then the probability of this mechanism will be $p' + \frac{1}{2^n} - (\frac{p'}{2^n})$

- Early registration: To prevent sending task information to non-relevant potential contractors, entities register their capabilities with the broker at the beginning. In this mechanism the operation that was *sharing* sensitive information was substituted with another operation that does not *share* the information. Hence, the privacy protection level of this mechanism becomes 1 for the given information.

- Terms and condition support: assuming there is p" probability that the agreement impacts the decisions of the participants of the agreement to not violate the agreement.

Given the above mechanisms:

$$PPL(PBCNP) = (1 - 2\varepsilon) * \left(p' + \frac{1}{2^n} - \left(\frac{p'}{2^n}\right)\right) * p"$$

4.6 Solution equivalency

Although the framework is capable of providing a privacy based interaction protocol, it is important to prove that the solutions reached using the traditional interaction protocol are still attainable using the new protocol.

Merging the operations of the protection mechanism with operations of the interaction protocol creates a new list of operations. Assuming S_1 is the solution that is achieved using IP. That indicates that the list of operations of the interaction protocol has been executed and completed. Assuming S_2 is the solution achieved by $PBIP$. Because the list of operations and protection mechanism are totally ordered in the list of operations in $PBIP$, the list of both interaction protocol and protection mechanism are executed and completed. If $S_1 \not\equiv S_2$, either 1) the list of operations in interaction protocol are not completed or 2) they are pre-empted by operations in protection mechanism.

The first condition cannot be true as the solution is achieved using $PB_IP$and therefore, all operations of the interaction protocol are executed and completed. However, the second condition can be valid. When the privacy protection mechanism operations are applied they might not be able to provide the necessary information for the interaction protocol operations to proceed. That can disrupt the sequence and another solution gets selected. Nonetheless, if the privacy protection mechanism operations cannot provide the information for the next interaction protocol operation, it is due to transferring sensitive information to outside of its exposure boundary. In the other word, if a solution is achievable through IP , it can be achieved using PB_IP as well, unless it does not meet the requirements for privacy. This is by design one of the objectives of the framework. The solutions that can result in privacy concern are not acceptable solutions. Therefore, any acceptable solution attainable by IP it can equivalently be reached by PB_IP as well.

4.7 Privacy Protection Management in the related works

In this section we provide a comparison between existing privacy models and practices and the proposed privacy framework. The proposed privacy protection framework also belongs to the architectural solutions, as it is an interaction-based framework and is

applied at the computational platform of CDS applications. In other words, it is not required to track the existing rules for each entity at each interaction. The interaction protocol assesses the level of PPL and risk of interactions, which enables decision-making processes at the interaction level. Therefore, this model exhibits characteristics of architectural approaches in privacy protection classes of solutions.

Many architectural based privacy solutions were proposed in the literature such as anonymization techniques [23], [12], [13], privacy utility trade off mechanisms, [5], [67], [14], social tradeoffs and proxy based privacy protection [16]. The proposed privacy protection framework does not consider any assumption in having trusted entities in the environment. In contrary many architectural-based solutions adhere to a particular setting of interactions [23], [13]. Furthermore, sensitivity of the aggregated information might be neglected in some forms of privacy models [23].

The utility tradeoff mechanisms are based on evaluating the information gain of the exchanged information [14], [5]. The work in [16] illustrates that these models do not necessarily reflect the preferences that each entity might have over their privacy. Information gain by itself does not convey the expectation of entities over privacy. In contrast, the proposed privacy model considers the sensitivity of information as a computation element measured by processing the exposure boundaries that are captured from each entity. In addition, the cost and risk evaluation functions that can be modeled by incorporation of various elements that represents the preferences of an entity can impact the decision making process.

The social tradeoff mechanisms measure the trust and intimacy relationship among entities as well as the information gain. However, these traits are very difficult to validate [16]. Typically, each entity has different criteria for intimacy and trust that might be included in cost and risk evaluation function that is applied by the entity. This enables the proposed model to be compatible with solutions that have the social aspects of interactions embedded inside of the privacy protection mechanism.

In proxy-based approaches, there is a mediator entity that acts as a proxy for other entities. It provides assessment over matching privacy expectation and the given privacy

protection in interactions. These approaches are based on the strong assumption that there are trusted entities that can be the proxy for other entities [16]. In contrary, the proposed model does not include assumptions over trusting entities. For every interaction, the exposure boundary of the *sharing* information is evaluated and the proper privacy protection mechanism is applied in the interaction protocol

4.8 Summary

Formal analysis of privacy is essential to apply privacy at the computation level. Accordingly a privacy model is proposed in the context of information management adequate for CDS environments that captures privacy as a mathematical object. The solution approach presented in this work is an architectural-based solution that is integrated at the interaction protocol. This work provides an analytical tool containing sets of formulated concepts that are essential for evaluating the state of privacy in computational systems. Additionally, it presents a framework for protecting privacy that is applicable on interaction of entities.

Modeling privacy in the context of information management, privacy in CDS can be reduced as the state of the exposure boundary of entities' states with the outside world. State of entities can be modeled by information. Therefore, when the information is communicated inside the boundary it is considered as non-sensitive and when it is exposed to outside of the boundary, it is sensitive information. The exposure can happen through *sharing* information or *disclosure* of information. Because entities can be classified as explicit and implicit information, the privacy concern is about the *disclosure* of sensitive implicit information. Any operation in implicit information that can extract sensitive information is considered to be unauthorized. There is an agreement among entities that the unauthorized operations are not executed on information. Disobeying the agreement and applying the unauthorized operation is considered as privacy violation. In contrast, utilizing a mechanism that can prevent or neutralize non authorized operations from execution is denoted as privacy protection.

The proposed model is a complete view for privacy in information management in CDS which is capable of reducing other types of privacy models to information management and therefore to the proposed privacy model.

In many applications of CDS environment, realizing the conditions for protecting privacy is incorporated with uncertainty. Entities have incomplete knowledge about communicated information as well as all operations in other entities. Therefore, it is required to provide metrics that can measure the uncertainty level towards privacy protection. The proposed privacy framework addresses the uncertainty of protecting privacy using the protection mechanism with the concept of PPL. PPL is the conditional probability of applying a protection mechanism under the space of an entity's information. This value becomes a measure for evaluating decisions in interaction of entities. Also the proposed framework provides formal analysis of applying the protection mechanism at the interaction level. The integration of the mechanism and interaction protocol results in privacy based interaction protocol that incorporates necessary, messages and sequences to support privacy protection at the interaction level.

Chapter 5

5 Privacy- Aware Agent Model and Implementation

The proposed privacy protection management framework is a generic approach to provide a privacy protection interaction protocol. It can be used as an analytical tool to identify concerns in an interaction protocol and can be incorporated with protection mechanisms. At the same time, it can be applied at the computation level and automate privacy protection management in the computation entity. Any solution achieved at the computation level requires problem solving and coordination with other entities. Thus far, we have proven that privacy resolution is essential at the computation in order to reach to acceptable solutions. Application of privacy protection management frameworks at the computation level provides privacy aware computation platform; all of the concepts of the privacy model are modeled by computational elements. Our contribution in this chapter includes designing and developing privacy-aware computation entity in agent-based model.

5.1 Privacy: Computation Concept in Computation Entity

Interactions are the mechanism of coordination used to resolve interdependency problem. Through this, computation entities can adequately be modeled as CIR agents in which they have knowledge, problem solving capabilities, interaction, and communication [19]. Figure 4 shows the logical architecture of a computation entity.

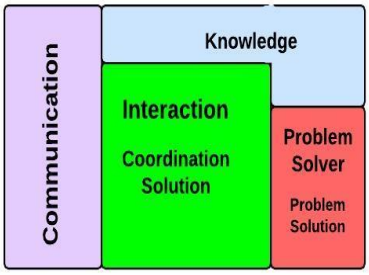

Figure 4. Logical Architecture of Computation Entity in CDS environments

$$Computation\ Entity\ \equiv < K_i, PS_i, In_i, Com_i >$$

In information management form of computation, entities are modeled as information and operation.

$$Information\ Management\ Entity\ e_i \equiv < I_i, O_i >$$

Knowledge in entities conveys all information regarding, intentions, believes and states of the entity. This includes the information regarding operations that the entity possesses and is capable of applying them. This allows modeling the knowledge as set of information and operations as the following:

$$K_i\ \equiv < I_i^k, O_i^k > , \subseteq (I_i\ , I_i^k\ ), \subseteq (O_i\ , O_i^k) \quad 17.$$

Problem solving in computation entities is an adjoined layer of the knowledge. It consists of operations to identify goals and required actions towards achieving it through the information acquired from knowledge. Because of this problem solving can be modeled as operation in information management.

$$PS_i\ \equiv\ O_i^{ps}\ , \subset (\ O_i^{ps}, O_i) \quad 18.$$

The computation entity at the interaction level encloses pattern of communication as well as decision-making on coordination to resolve interdependency problem. Interaction layer is adjacent to the knowledge, problem solver and communication layers. Through this, the interaction can be modeled as information and operations

$$In_i\ \equiv < I_i^{In}, O_i^{In} > , \subset (I_i^{in}, I_i), \subset (O_i^{in}, O_i) \quad 19.$$

The communication layer encompasses the messages that will be communicated to other entities, but it does not interfere with coordinating the decision-making processes. The communication layer is modeled as information in information management.

$$Com_i\ \equiv\ I_i^{com}\ , \subset (I_i^{com}, I_i) \quad 20.$$

Privacy is the concern of decentralized environments in which entities *share* information through communication-based interactions. Naturally, privacy as a computation concept is inherently expressed at the interaction level. Privacy solutions at the entity level facilitate entities by interacting with other entities that are driven by privacy aware interactions. The solution at the computation entity allows the environment to achieve global solutions in which entities' privacy is respected. Figure 5 depicts the relationship of the privacy solution with other layers in the computation entity.

Applying proposed privacy protection management frameworks at the computation level incorporates privacy protection management directly with the interaction. Privacy solutions stand between communications and interactions to consolidate interactions with privacy-based interaction protocol and privacy protection management.

As described in the privacy protection management framework, interaction protocols can be modeled as sets of messages and sequences thereof:

$$IP \equiv < M, S_M >$$

$$M \equiv \{m_1, \dots, m_m\}, 1 \leq m \leq Z$$

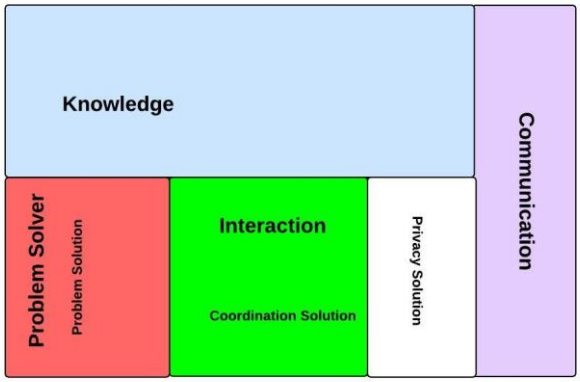

Figure 5. Privacy solution in relation with interaction in the computation entity in CDS environment

Each message in the interaction protocol conveys content; this content involves a a sending and receiving entity and operations that transfer the message.

$$m_m \equiv < e_s, e_r, C_m, O_{s,m} > , e_r: Receiver, e_s: Sender, C_m: \ content \ , \in (C_m, I_i), \ \in (O_{s,m}, O_{s)} \ 21.$$

Sequences are constructed by patterns of exchanging messages.

Assuming,

$$M^k: All\ sequences\ in\ M\ with\ k\ length$$

Then

$$M^* \equiv \bigcup_{k=1}^{Z} M^k \ , M^*: All\ possible\ sequqnces\ given\ the\ set\ M$$

Therefore,

$$S_M \ \subset \ M^*$$

$$S_M = [s_1, \dots, s_q], 1 \leq q \leq Q \quad 22.$$

Each sequence carries multiple messages

$$s_q = [m_{a+1}, \dots, m_{a+X}], 0 \leq a \leq V, 1 \leq X \leq Z \ \ 23.$$

which can include several sub-sequences:

$$ss_{q,t} \equiv [m_{a+l}, \dots, m_{a+p}], 1 \leq l, p \leq X, 1 \leq a \leq V,$$
$$1 \leq t \leq T, ss_{q,t}: Subsequent\ of a\ sequence \quad 24.$$

Let s_q^* be the set of all subsequences of a sequence. Then:

$$s_q^* \equiv \bigcup_{1 \leq t \leq Z} ss_{q,t}$$

25.

As messages are bound to operations that deliver them, $ss_{q,t}^{o}$ represents all of the operations of a subsequence:

$$ss_{q,t}^{o} \equiv \bigcup_{h=l}^{p} o | o \equiv O_{i,a+h} \ \wedge \in (m_{a+h}, ss_{q,t})$$

26.

$$ss_{q,t}^{o} \equiv \left[o_{i,a+l}, \dots, o_{i,a+p}\right], 1 \leq l, p \leq X, 1 \leq a \leq V \qquad 27.$$

Therefore, the execution of a the operations of a subsequence on the set of messages of an interaction protocol is denoted as

$$\bar{\bar{ss}}_{q,t}^{o}(M) \equiv \ \bar{\bar{o}}_{i,a+l}(C_{a+l}, \bar{\bar{o}}_{i,a+(l+1)}(\dots, \bar{\bar{o}}_{i,a+p}(C_{a+p})) \quad 28.$$

5.1.1 Privacy Protection Management

To capture privacy at the computation level and provide protection mechanism, it is required to incorporate privacy in interactions. Interactions are steered by interaction protocols that can be modeled as messages and sequences of messages. Privacy Protection Management is responsible in identifying privacy concerns in interaction protocols and providing privacy based interaction protocol that encompasses the protection operations to protect privacy. The logical architecture of the privacy protection management layer in the computation entity in CDS environments is depicted in Figure 6.

5.1.2 Capturing information and the exposure boundaries

Privacy in the context of information management is the state of exposure boundary of information that includes entities for which *sharing* information can happen. Knowledge in the computation entity includes all information, intentions, believes as well as the exposure boundary of information

$$\subseteq (I_i, I_i^k) \ , \forall \ k, \subset (E_{i,k}, I_i^k)$$

Information is *shared* through messages of interaction protocol. By capturing the receiver entities of the messages in the interaction protocol, the participating entities in the interaction will be identified. Based on equation 21:

$$m_m \equiv < e_s, e_m, C_m, O_{s,m} > , e_s : Receiver , e_m : Receiver, C_m : \ content \ , \in (C_m, I_i) , \\ \in (O_{s,m}, O_s)$$

Therefore:

$$R^* \equiv \bigcup_{m=1,s=1}^{Z} e_m, e_s | \in (< e_s, e_m, C_m, O_{s,m} >, M), R^* : \ Participating \ Entities$$

29.

Figure 6 shows the logical architecture of the privacy aware computation entity. Within this architecture, the Exposure Boundary layer collects the exposure boundaries of the information that is *shared* in interaction protocols.

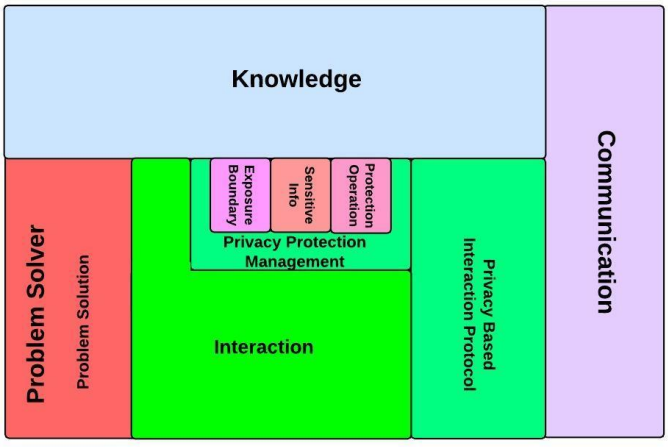

Figure 6. The logical architecture of privacy protection management in computation entity in CDS environments

5.1.3 Identifying the Sensitive Information

By applying the framework principles and given the exposure boundaries, the sensitive information can be captured as the following:

$$I_i^s \equiv \bigcup_{k=1,j=1}^{k\leq N,j\leq W} \left(I_{i,k}, e_j\right)| \in (e_j, (R^* - E_{i,k})), I_i^s : all\ sensitive\ Information$$

30.

As depicted in Figure 6, the layer of "Sensitive Information" is adjacent to exposure boundary and interaction. This allows this layer to capture the necessary elements from the exposure boundary and interaction protocol to identify sensitive information.

5.1.4 Diagnosing Privacy Concerns in the Interaction Protocol

Interaction protocol follows a sequence of messages among entities. These messages have content that carries required information to follow the protocol. In this context, messages are tied to operations that deliver the content from one entity to another. This positions messages in conjunction with operations equivalent to *sharing*.

Considering equation 21:

$$M \equiv \{m_1, \dots, m_m\}, 1 \leq m \leq Z$$

$$m_m \equiv < e_s, e_m, C_m, O_{s,m} >, e_s : Receiver, e_m : Receiver, C_m :\ content\ , \in (C_m, I_i), \in (O_{s,m}, O_s)$$

$O_{s,m}$ delivers the messages transferred to the communication layer. Therefore:

$$O_{s,m}\ \equiv = \left(I_m^k, \bigcup (C_m, I_m^k)\right)$$

Based on equation 5 in the privacy protection management framework, *Sharing* C_m is :

$$S(C_m, e_m) \equiv = \left(I_m, \bigcup (C_m, I_m)\right)$$

We also know that:

$$\subseteq (I_m, I_m^k)$$

This shows that:

$$S(C_m, e_m) \equiv = \left(I_m^k, \bigcup (C_m, I_m^k)\right)$$

Therefore:

$$S(C_m, e_m) \equiv \ O_{s,m}$$

The content of messages convey information that might disclose sensitive information in conjunction with other messages of the sequence. This results in *disclosing* sensitive information when the sequences of messages are exchanged. The sensitive information is computed by capturing the exposure boundaries. Therefore, evaluating sub-sequences of the interaction protocol to identify *disclosing sensitive* information is essential.

$$H_i^* \equiv \bigcup_{q=1,j=1,t=1,k=1}^{q\leq Q, j\leq W, t\leq T, k\leq N} (ss_{q,t}, I_{i,k}) | \ (= (\bar{\bar{ss}}_{q,t}^{o}(M), I_{i,k}) \ \wedge \ I^s\left(I_{i,k}, e_j\right))$$

31.

Also, non-sensitive information can be used as auxiliary information to transform implicit sensitive information to explicit. Typically, the preventive mechanisms cannot be applied for auxiliary information as they are *shared* within the exposure boundary. To deal with this, entities comply with agreements and applying punishing mechanisms. The concern regarding the auxiliary information can be identified through exploring the receivers and the information *shared* with them in a sequence of messages in the interaction protocol.

$$A_i^* \equiv \bigcup_{q=1,j=1,t=1,k=1}^{q\leq Q,j\leq W,t\leq T,k\leq N} ([o], I_{i,k})| \in (o, ss_{q,t}^{o}) \wedge \bar{\bar{o}} \equiv S(I_{i,k}, e_j)$$

32.

Hence, the concern points in the interaction protocol can be identified as follows:

$$D_i^* \equiv \bigcup (H_i^*, A_i^*), D_i^*: Concern\ points$$

33.

5.1.5 Determining Required Protection Operations with adequate PPL

Protection operations are part of the knowledge of the entity. Entities can utilize various protection operations that are registered within the knowledge of the entity. Protection mechanisms such as differential privacy anonymization [62], private bid-communication [57] and contractual operation execution[68] are examples of protection operations that can dynamically be registered in an entity and be applied on the interaction protocol. Each protection operation comes with the associated PPL that will be used as a measure to evaluate the privacy state of the privacy-based interaction protocol.

Protection Operation Registration:

$$Reg(e_i, \mu, PPL) \equiv = \left(O_i^k, \bigcup (O_i^k, \mu)\right) \wedge = \left(I_i^k, \bigcup (I_i^k, PPL(\mu))\right)$$

A protection operation in a computation entity when the $Requested\ PPL\ =$ A is denoted as:

Protection Operation:

$$\mu_{i,k} \equiv \mu \mid \in (\mu, O_i^k) \wedge \exists j, \in ((i_{i,k}, e_j), I_i^s) \wedge \exists \subset (S, I_i) \wedge \in (I_{i,k}, S) \wedge PP(\mu|S) \wedge PPL(\mu) > A$$

Depending on the sequence of messages in the interaction protocol and the identified privacy concern, an adequate protection operation is required to be applied. Based on the above definition, Protection Operation layer performs analysis on the given sequences and the expected PPL value to retrieve the adequate protection operation among available protection operations.

5.1.6 Expanding the Messages and Sequences

The privacy protection management framework introduces three forms of expansions in interaction protocol: prefixing, appending and generic. In the generic forms of expansion, for each of subsequences tagged as concern point, the privacy-based sequence will be substituted with the original one. The concerns marked as auxiliary will be extended with punishing protection operations as well as the structure to include the adequate agreements.

Depending on the content that is *shared* in a subsequence, it might be tagged as concern point multiple times. Let $\lambda_{q,t}$ be the sensitive information that $ss_{q,t}$ is tagged for. Based on equation 24 and 31:

$$\lambda_{q,t} \equiv \bigcup_{k=1}^{k\leq N} I_{i,k} | \in ((ss_{q,t}, I_{i,k}), H_i^*)$$

$$\lambda_{q,t} \equiv \{I_{i,k}, \dots, I_{i,u}\}$$

Then we can retrieve the protection record of the information in $\lambda_{q,t}$

$$\mu(\lambda_{q,t}) \equiv \bigcup_{k=1}^{k\leq N} \mu_{i,k} | \in (I_{i,k}, \lambda_{q,t})$$

and let $\lambda'_{q,t}$ be the information that might be used as auxiliary information in a sequence:

$$\lambda'_{q,t} \equiv \bigcup_{k=1}^{k\leq N} I_{i,k} | (ss_{q,t}, I_{i,k}) \in A_i^*$$

$$\lambda'_{q,t} \equiv \{I_{i,k'}, \dots, I_{i,u'}\}$$

Similarly the protection record of the information in $\lambda'_{q,t}$ is

$$\mu(\lambda'_{q,t}) \equiv \bigcup_{k=1}^{k \leq N} \mu_{i,k} | I_{i,k} \in \lambda'_{q,t}$$

Also

$$ss^{o}_{q,t} \equiv [o_{i,a+l}, \dots, o_{i,a+p}], 1 \leq l, p \leq X, 1 \leq a \leq V \quad 34.$$

Then

$$PB\_Seq(ss_{q,t}) \equiv Pun(Prev\left(ss^{o}_{q,t}, \mu(\lambda_{q,t})\right), \mu(\lambda'_{q,t})) \quad 35.$$

$$Prev\left(ss^{o}_{q,t}, \mu(\lambda_{q,t})\right) \equiv Prefxing(ss^{o}_{q,t}, \mu(\lambda_{q,t})) \quad 36.$$

$$Pun\left(ss^{o}_{q,t}, \mu(\lambda'_{q,t})\right) \equiv$$

$$Prefxing\Big([< Agreement, Reject >, <$$

$$Agreement, Confirm >], Appending\left(ss^{o}_{q,t}, \mu(\lambda'_{q,t})\right)\Big) \quad 37.$$

The privacy based sequence for the subsequences that do not belong to concern points will stay as the original subsequence. This is due to $\mu(\lambda_{q,t}) \equiv \emptyset$ and $\mu(\lambda'_{q,t}) \equiv \emptyset$.

The sequences of the privacy-based interaction protocol are the set of all sequences or their privacy based sequences substitutions if they are among the concern points.

$$PB\_S_M \equiv \bigcup_{q=1,t=1}^{q \leq Q, t \leq T} PB\_Seq(ss_{q,t})$$

38.

The messages that are exchanged in these sequences will form the set of messages that the privacy-based interaction protocol utilizes.

$$PB_M \equiv \bigcup_{q=1,t=1,s=1,r=1}^{q\leq Q,t\leq T,s\leq R,r\leq R} < e_s, e_r, C_m, O_{s,m} > \left| \in (O_{s,m}, ss_{q,t}^{o}), \subset (C_m, I_i), \in (e_r, R^*), \right. \in (e_s, R^*)$$

39.

This completes the necessary elements to present the privacy-based interaction protocol.

$$PB\_IP \equiv < PB_M, PB\_S_M >$$

There will be more discussions on Contract Net Protocol (CNP) as an example of interaction protocol in Chapter 6. This protocol is converted to the above elements and the computation entity adopted the privacy based contract net protocol.

5.1.7 Expanding Computation Entity with Privacy Solution

Earlier in this chapter we provided the analysis on application of the proposed privacy protection framework at the computation level. The elements of the privacy protection management are determined using computational concepts at the entity level. In this section we show that privacy as a computation concept in a computation solution.

In distributed decentralized computing systems, the solution is achieved by capturing the solution at the problems solver PS and the coordination solution (CS) in interaction [19].

$$S \equiv f(PS, CS)$$

Consider the following example:

$Agent\ 1: Is\ within\ the\ exposure\ baoundary\ of\ I_{i,k}$

$Agent\ 2: Requires\ capabilities\ of\ another\ agent\ outside\ of\ exposure\ boundary$

Agent 1 and Agent 2 can interact with each other and share $I_{i,k}$ as they belong to the exposure boundary. However, because Agent 2 requires interacting with another agent which is outside of the exposure boundary, $I_{i,k}$ will be disclosed to entities outside of the exposure boundary and they can extract it. Hence, the solution is not feasible and the solution does not exist (Figure 7).

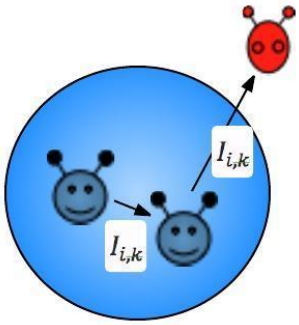

Figure 7. Solution without privacy protection does not exist

If privacy solution is not applied the system can reach to a feasible solution. However, by applying privacy protection mechanism, the solution can be reached. For instance the Figure 8 shows a solution that can be approved using privacy protection mechanism applied on information before it goes outside of the exposure boundary.

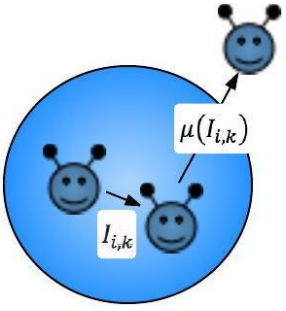

Figure 8. Solution Exists with Applying Privacy Protection Mechanism

When the computation solution reaches to a solution, it is essential that the solution can provide an adequate level of privacy protection. This indicates that the coordination solution compares the possible choices in regards to privacy for instance:

$$S \equiv f(PS, CS) in\ relation\ to\ e_j?$$

$$S \equiv f(PS, CS) in\ relation\ to\ e_j\ when\ sensitive\ information\ is\ disclosed?$$

This illustrates that the computation function is expecting a new dimension to be able to make a decision on the solution. There could be possible solutions that can perform the requested task, but the one with privacy will be accepted. This expands the computation with a new parameter that reflects the solution for privacy:

$$S \equiv f(PS, CS, PrS)$$

5.2 Implementation Challenges

The sequences of the privacy-based interaction protocol are the set of all sequences or their privacy based sequences substitutions if they are among the concern points in equation 39.

$$PB\_S_M \equiv \bigcup_{q=1,t=1}^{q\leq Q,t\leq T} PB\_Seq\left(ss_{q,t}\right)$$

Based on equation 39, the messages that are exchanged in these sequences will form the set of messages that the privacy-based interaction protocol utilizes.

$$PB_M \equiv \bigcup_{q=1,t=1,s=1,r=1}^{q\leq Q,t\leq T,s\leq R,r\leq R} < e_s, e_r, C_m, O_{s,m} > \left| \in (O_{s,m}, ss_{q,t}^{o}), \subset (C_m, I_i), \in (e_r, R^*), \in (e_s, R^*) \right.$$

This completes the necessary elements to present the privacy-based interaction protocol.

$$PB\_IP \equiv < PB_M, PB\_S_M >$$

The component diagram of the privacy protection management is presented in Figure 9.

5.3 JIAC: Implementation Platform

JIAC (Java Intelligent Agent Component) is a framework for developing distributed heterogeneous, complex systems. This platform supports the developments of Multi-Agent Systems (MAS) through features such as [28]:

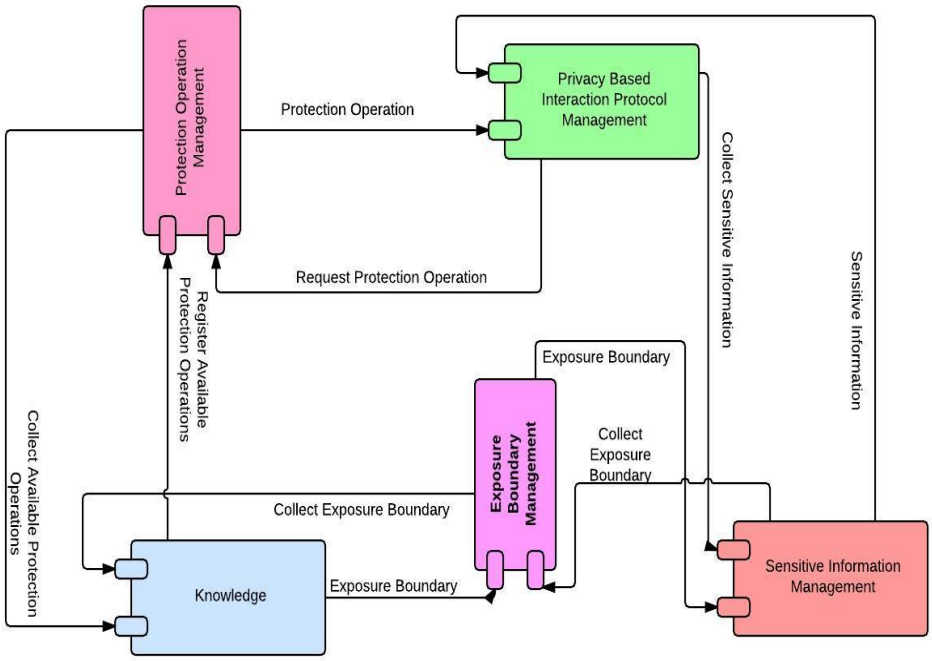

Figure 9. The Competent architecture of Privacy Protection Management in Computation entity

- Spring-Based Component System
- ActiveMQ-based messaging
- JMX-based management
- Transparent distribution

Applying the privacy protection management framework at the computation level, requires expanding on the computation entity. It requires developing a CDS environment

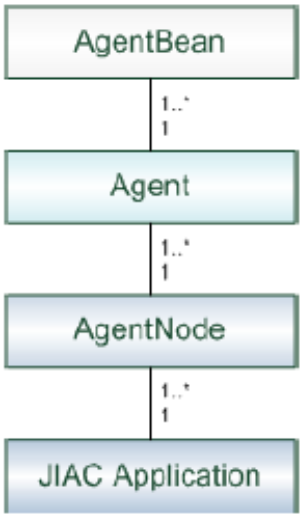

Figure 10. JIAC Applications and the relationships to other JIAC Concepts [28]

where autonomous self-interested entities interact. JIAC platform enables developing distributed decentralized setting in which the communication and agent life cycle management is steered by the platform. JIAC flexibly incorporates new behaviors to the agent which is essential for expanding new message types at the interaction protocol. On that account, JIAC is an adequate platform for implementing the privacy aware computation entity.

5.3.1 JIAC Platform

JIAC applications typically inherit decentralized distributed context, which consist of multiple Agent Nodes. The AgentNode is a computation-service platform that is architected as distributed layer providing services to agents. Each AgentNode includes several Agent Components. Each of which contains classes of beans that specify the behavior of the Agent (Figure 10).

5.3.2 Agent Life Cycle

The agent life cycle refers to the state that an agent can be in which is steered by the AgentNode . Additionally the agent and agent beans can perform processes when the state changes in an agent. Figure 11 shows the agent life cycle in JIAC platform.

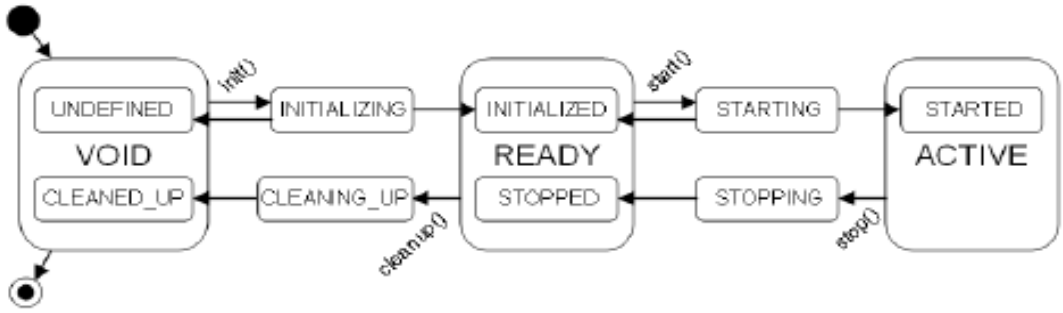

Figure 11. Agent Life Cycle in JIAC Platform [28]

5.3.3 Agent Actions

Actions are part of the trait of agent beans in JIAC that allows the asynchronous execution of behaviors in agents. All the operations including the protection operations and interaction operations are implemented as an action in the JIAC platform. Actions can be added dynamically to the memory of the agent. It can scale up to the agent node as well as direct the agent to be accessible by search inquiries. In this work, we have introduced the actions at node level. Actions are searched by template specification which specifies that characteristics of an action to be called. When the protection operations are registered, the template of their action is added to the agent and it will be called when the action is searched through the agent memory. Actions can perform send operations as well as performing processing operations. The flexibility of the dynamic action allowed us to implement the operations of the sequences of the interaction protocol as an action within an agent. Before the agent gets to the ready state, the action list is updated so that the agent accesses the necessary actions.

Figure 12 shows part of the execution of the bean that dynamically adds the template of the action to the agent.

```
@Override
public List<? extends IActionDescription> getActions() {
    List<Action> ret = new ArrayList<Action>();

    Class<?> c=null;
    Object obj=null;
    try {
        c = Class.forName(class_name);
        obj=c.newInstance();
    } catch (ClassNotFoundException e) {
        e.printStackTrace();
    } catch (InstantiationException e) {
        e.printStackTrace();
    } catch (IllegalAccessException e) {
        e.printStackTrace();
    }
    Action echo = new Action(action_name, (IEffector) obj, new Class[]{String.class,String.class}, new Class[]{})
    echo.setScope(ActionScope.NODE);
    ret.add(echo);
    return ret;
}
```

Figure 12. Adding Dynamic Action to the agent at Node level

5.3.4 Privacy Protection Management in JIAC Agent

The following is the details of implementation of privacy aware computation entity within JIAC agent platform. Some of the proposed components are employed to resolve the requirements and restriction of implementation platform.

In JIAC application, interaction sequences can be modeled as set of actions that performed by entities. Each of the messages and protection operations is captured as actions.

The Privacy Protection Management expands the interaction protocol with adequate protection operations and provides privacy based interaction protocol. Privacy Based Interaction Protocol manages all interactions of the computation entity with others through which the adequate privacy protection operations are applied. The messages and sequences of messages that are sent for communication are managed by the privacy aware interaction protocol. The functionalities of this layer can be categorized as follows:

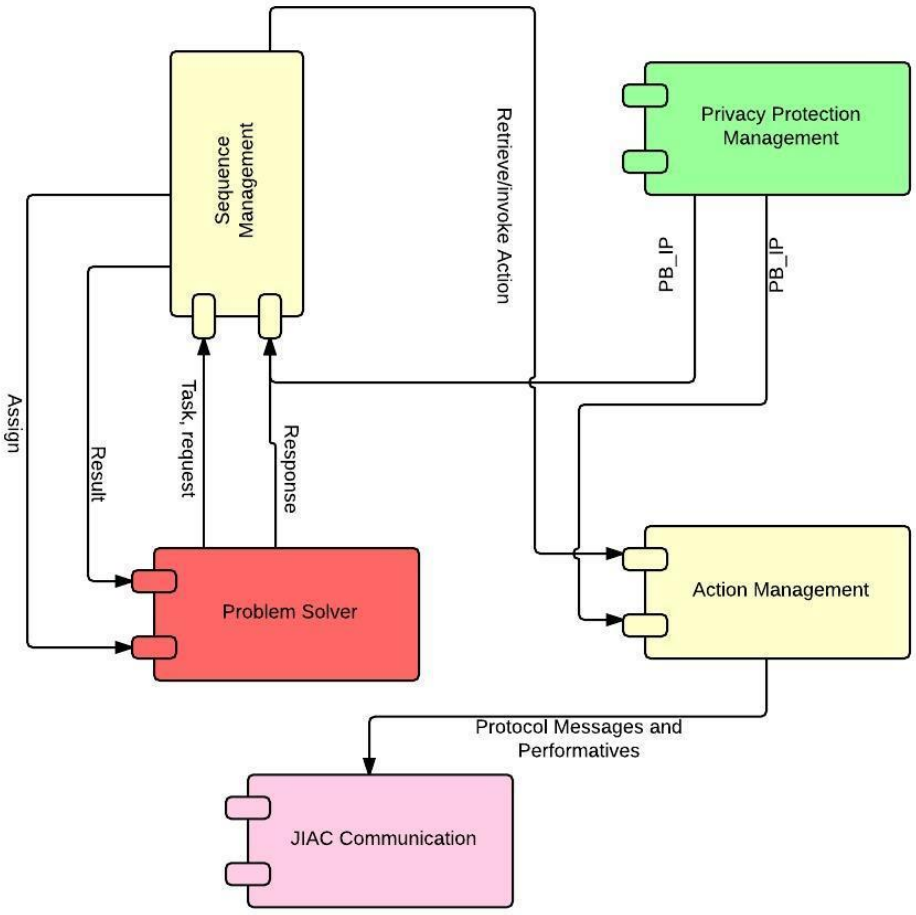

Figure 13. Component diagram of the implemented JIAC agent

A. Expanding the message indexes with the new and modified message types.
B. Managing sequences of the protocol upon receiving new messages when entities interact.

A is packaged as the functionalities of "Action_Management" component in the component architecture diagram. B is the functionality of the "Sequence_Manegemt" component in the component diagram. The component architecture is shown in Figure 13. The class diagram of the Privacy Protection Management component is provided in Figure 14. This component applies the privacy protection management framework at the computation entity and expands the interaction protocol with the privacy based interaction protocol. Also, the class diagram of the components of the JIAC agent is provided in Figure 15.

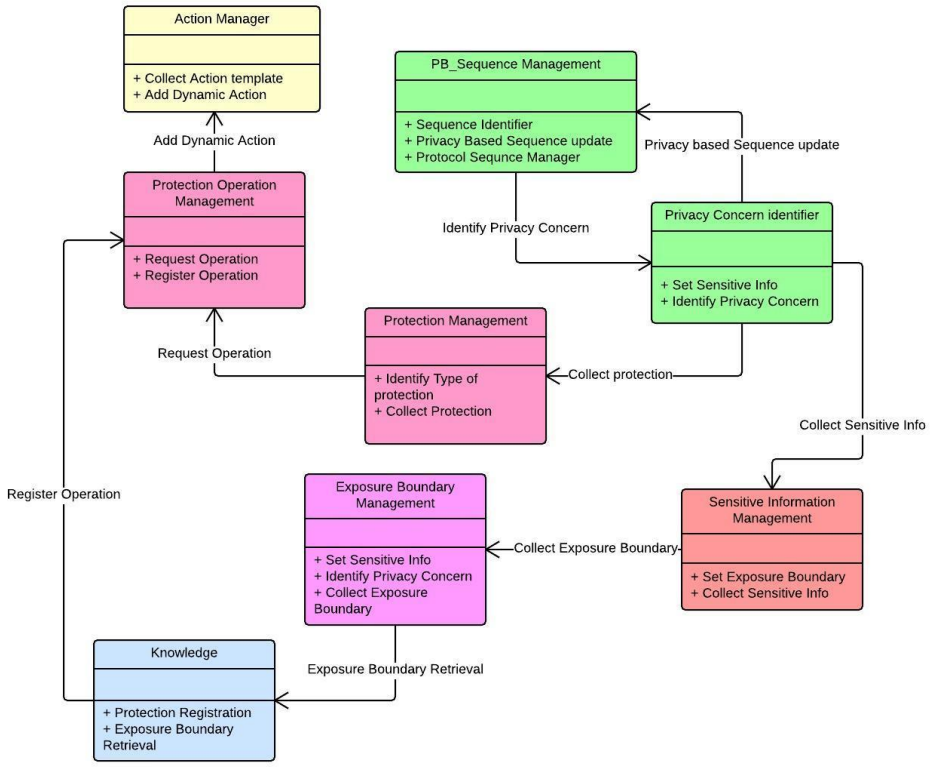

Figure 14. Privacy Protection Management component

5.4 Summary

The proposed privacy protection framework is a generic model that can be used as an analytical tool for identifying privacy states of interaction protocol as well as getting applied in contexts such as computation level. Capturing privacy as a computation concept necessitates incorporating privacy in the computation entity at interaction level. The computation entity in CDS environment requires resolving interdependency problem through interaction. The privacy based interaction protocol enables the entity to become privacy aware in its interactions. In this chapter, we provided the computational aspect related to privacy protection management framework. We also adopted the JIAC agent component ware as the implementation platform for the entities that are implemented as agents. Every operation in JIAC agents is modeled through actions. The messages and sequences of the privacy based interaction protocol are incorporated in actions each of which is dynamically added to the agent memory. In this chapter, the supporting argument to validate the sufficiency and adequacy of the proposed privacy model and

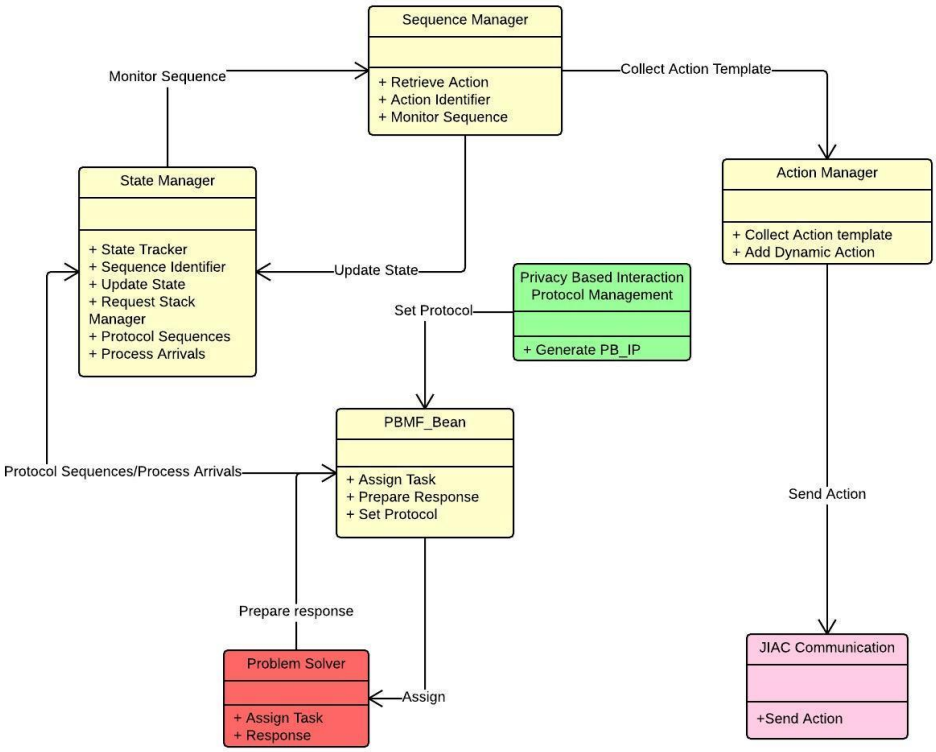

Figure 15. Class Diagram of components of JIAC Agent

privacy protection management in CDS is presented. The proposed privacy protection framework that is applied at the computation level expands the computation solution with new parameter that reflects the coordination with entities and performing actions that can protect privacy.

Chapter 6

6 A Privacy-based Interaction for CNP Protocol

Interaction protocols are the mechanism used to resolve interdependency problem in CDS. One of the approaches to interact is through negotiation [69], [70]. Contract Net Protocol (CNP) is a negotiation-based protocol that is applied for task allocation in CDS [71], [69]. Because of capability interdependency among various entities of CDS, they assign their tasks to others. CNP is an assignment interaction protocol that initially was proposed for distributed problem solving among various sensors. The messages of this protocol convey information that might disclose sensitive information. By applying the

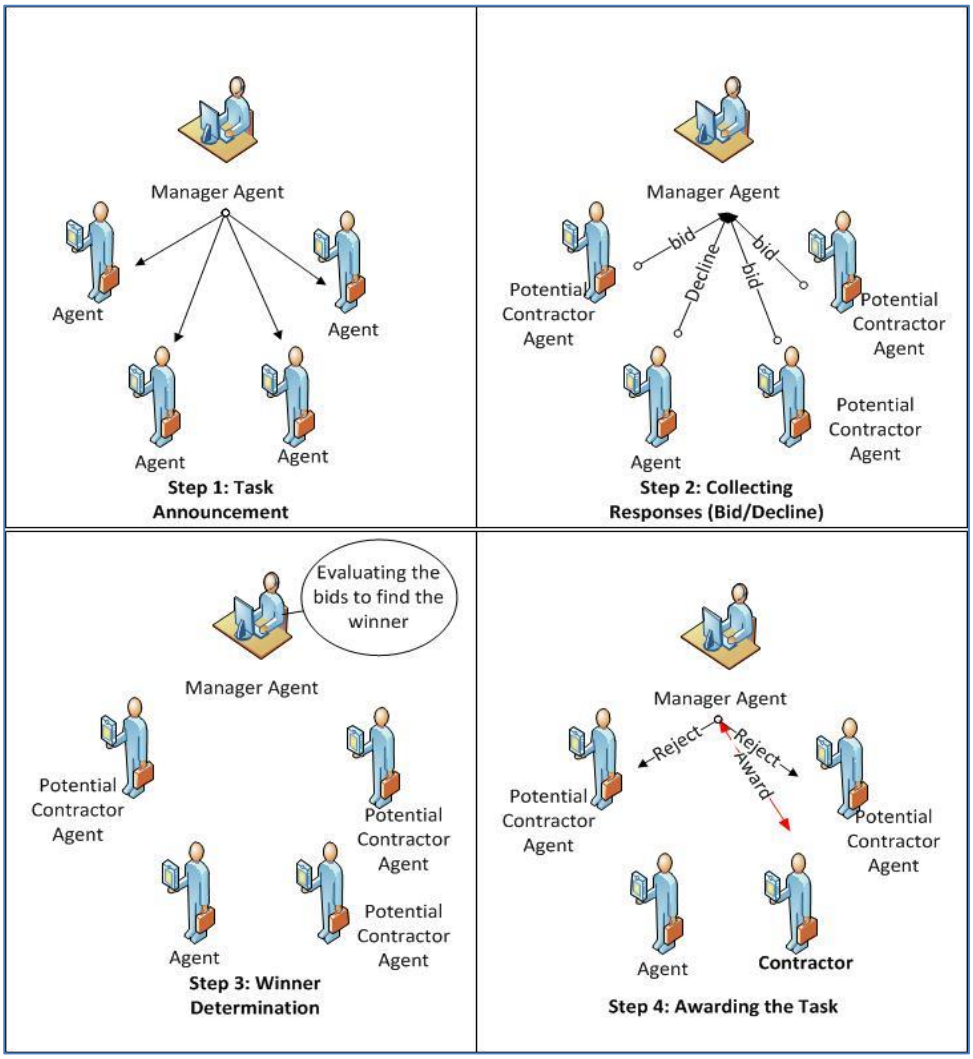

Figure 16. Contract Net Protocol

privacy protection management framework on this protocol, the privacy concerns are identified and proper protection operation is applied.

6.1 Contract Net Protocol

CNP contains manager sensors that announce a task to other sensors. The rest of entities [potential contractors] that are capable of executing the task compete for acquiring it. The entity that its proposal is accepted becomes the contractor and delivers the result after executing it [71]. Figure 16, shows different stages of contract net protocol to resolve an assignment interdependency. CNP can be expanded to be applied on brokering architecture in which to resolve the capability interdependency in such setting, entities adopt CNP as the interaction protocol where the brokering layer acts as the manager. The sequence presentation of CNP as discussed in previous section includes a set of sequences for operations in task announcement, winner determination and task execution.

$$\text{CNP} = \{< Request(TaskAnnouncement), Task\_Announcement(), Propose(Bid),$$

$$\text{Accept\_Proposal}, \text{Inform}, \text{Response} >,$$

$$< Tas\_Announcement, Propose(Bid), Reject\_Proposal >,$$

$$< Task\_Announcement(task), Reject >\}$$

“Task announcement” is a phase in CNP to inform entities about the characteristics of a task. Manager is responsible to send the “task announcement” to other entities. The information embedded in “task announcement” contains: eligibility specification, task abstraction, bid specification and expiration time. “Eligibility specification” conveys a list of criteria that an entity needs to have to be eligible to submit a bid. “task abstraction” includes information that briefly explains the task to be executed. “bid specification” is an indicator for potential contractors to know how the manager wants to receive the bids. Finally, the “expiration time” is a deadline for the execution of the task. After awarding the task to the contractor, they still can interact by information messages. These messages can be the interim or final report of the execution of the task. They also can be “REQUEST” messages. If the contractor needs to receive more information to complete

the task, it will send a “REQUEST” message to the manager. If the requested information is not in the “MUST HAVE” list and it is transferable, the manager collects the information and sends it to the contractor. If the contract requires other capabilities from others to execute the task, it can create sub tasks out of the executing task and request for help from other entities. The manager can specify the entities that the contractor can send the subtask to [71].

“Collecting responses” or bid proposal is a process that all entities that receive the task announcement evaluate the task and if they are capable of executing the task, they send their bids for the task, otherwise; they reject it.

In winner determination process, the manager collects all the bids from the potential

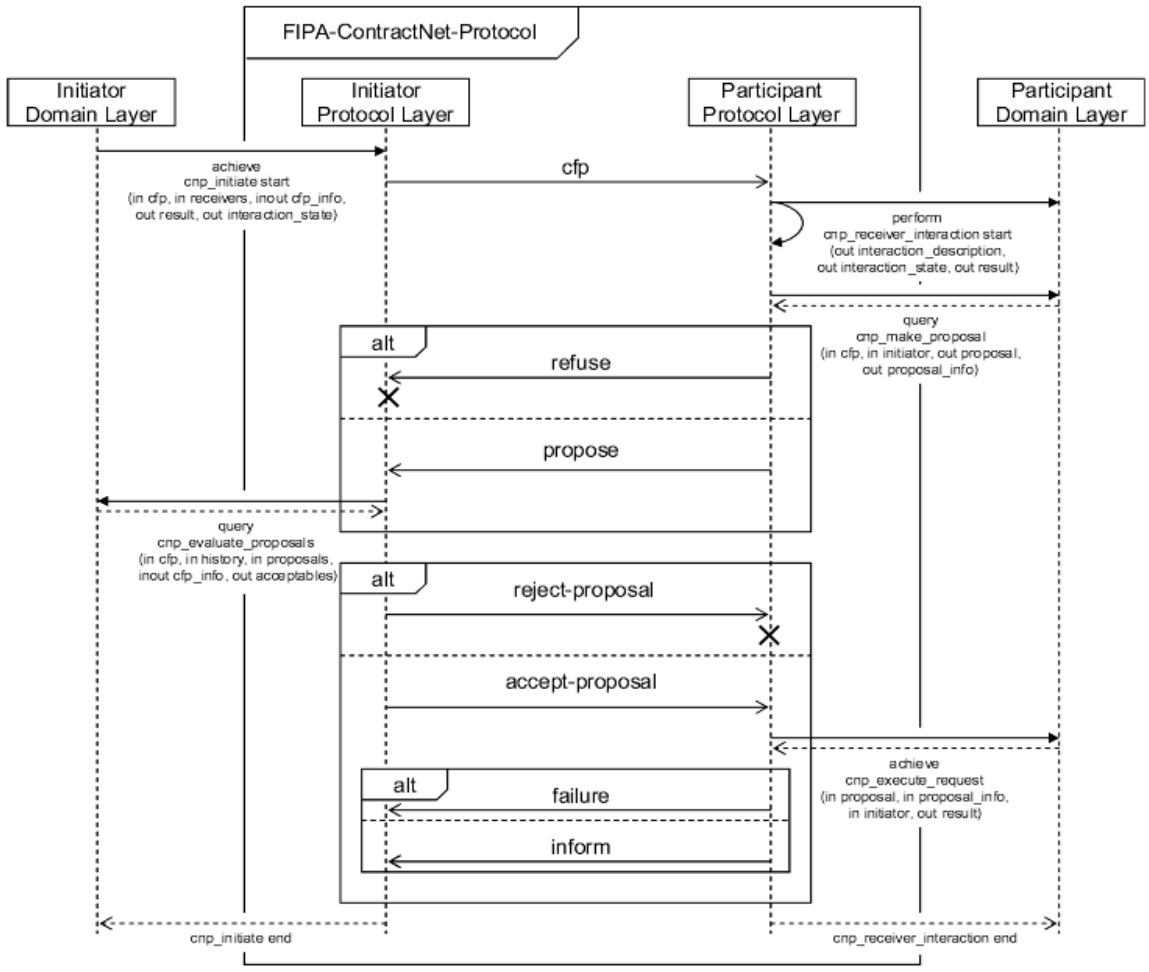

Figure 17. Traditional CNP

contractors and selects the best bid. The selected contracted will be awarded in the awarding process. The sequence diagram of the traditional CNP is shown in Figure 17.

Entities *share* information using the message types and sequences of messages in CNP. However, privacy concerns in *sharing* information in CNP is not considered and providing privacy protection at this protocol is lacking. In this protocol,

- Requester (e_r): The entity that has a task and needs a manager entity to find a contractor for executing the task
- Manager (e_m): the entity that searches for the contractor that can perform a task
- Potential Contractor (e_p): the entity that is a candidate for being awarded by the task
- Non eligible potential contractors (e_n): the entity that is not capable of executing the task
- Sub contractors (e_s): the entity that is awarded by a subtask
- Contractor (e_c): the entity that is awarded by the task

Given the information that is *shared* in CNP, here are the exposure boundaries associated to the information:

$I_{r,t} = \{name, task\_announcment\}$ and exposure boundary $I_{r,t}$ is $E_{r,t} = \{manager\}$

$I_{p,b} = \{\text{bid}\}$ and the exposure boundary $I_{p,b}$ is $E_{p,b} = \emptyset$

$I_{r,j} = \{result\_description\}$ and the exposure boundary of $I_{r,j}$ is $E_{r,j} = \{Contractor\}$

$I_{r,h} = \{\text{task\_history}(e_r)\}$ and the exposure boundary of $I_{r,h}$ is $E_{r,h} = \emptyset$

$I_{r,s} = \{\text{result\_history}(e_r)\}$ and the exposure boundary of $I_{r,j}$ is $E_{r,s} = \emptyset$

$I_{r,g} = \{\text{subtask}(I_{r,t})\}$ and the exposure boundary of $I_{r,g}$ is $E_{r,g} = \{e_v, \dots, e_f\}$

CNP has been used in many examples of CDS environments. However, this protocol does not apply privacy protection mechanisms while *sharing* information among entities. Using the proposed privacy protection management framework, we can transform CNP to a privacy based contract net interaction protocol. The privacy protection management framework considers the exposure boundaries and identifies the sensitive information is

shared or may be implicitly disclosed to other entities. Accordingly, here is the sensitive information that is identified in CNP for which privacy protection mechanism should be applied:

- $\mathrm{I}^{\mathrm{S}}(I_{r,t}, e_p)$
- $\mathrm{I}^{\mathrm{S}}(I_{p,b}, e_m)$
- $\mathrm{I}^{\mathrm{S}}(I_{r,j}, e_m)$
- $\mathrm{I}^{\mathrm{S}}(I_{r,t}, e_s)$
- $\mathrm{I}^{\mathrm{S}}(I_{r,h}, e_m)$
- $\mathrm{I}^{\mathrm{S}}(I_{r,t}, e_n)$
- $\mathrm{I}^{\mathrm{S}}(I_{r,s}, e_c)$

6.2 Sensitivity Analysis

$E_{r,t}$ which is the exposure boundary of $I_{r,t}$ only includes the manager entity. $I_{r,t}$ is the combination of identity of the requester and the task that is submitted to the manager. The structure of task announcement is as the following:

<task-announcement> ⇒ TASK-ANNOUNCEMENT [name] {task-abstraction} {eligibility-specification} {bid-specification} [expiration-time]

Because the exposure boundary of $I_{r,t}$ does not include potential contractors, combination of task announcement and identification of entities in CDS is sensitive information in relation with potential contractors. For instance, the task announcement of $\mathrm{e_r}$ might include the inquiry to activate light monitoring service in particular time slots of days. As an example the building management software that tracks various buildings of a house holding company is using light sensors for security reasons but they only activate them in low human traffic hours. They also work with entities monitoring power consumptions. These sensors might be provided by other parties, though they belong to the same environment. The task announcement and the identity of the requester can implicitly refer to hours of low traffic and sleep time $I_{r,k}$. This information becomes sensitive in relation with other third parties.

$$\exists\ o_{p,w} |\ \ o_{p,w}(I_{r,t}) \equiv I_{r,k} \wedge \mathrm{I}^{\mathrm{S}}(I_{r,k}, e_p)$$

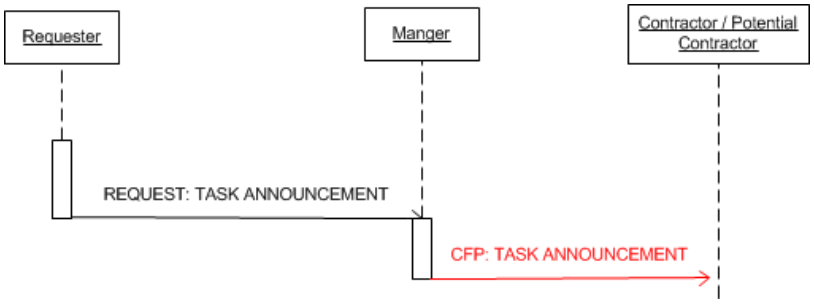

Figure 18. Task announcement is sent to all potential contractors

In traditional CNP, not differentiating potential contractors and *sharing* the task announcement discloses sensitive information (Figure 18).

$$\mathrm{S}(I_{r,t}, \mathrm{e_p}) \rightarrow \mathrm{D}(I_{r,k}, \mathrm{e_p})$$

The biding structure in CNP includes the identity and node abstraction which include the specification of the node that is providing the proposal. Node abstraction also includes the information that the contractor might need in case of being a winner.

<bid> ⇒ BID [name] {node-abstraction}

When entities want to compete with each other over the task, they submit bidding information with their identity. The exposure boundary of potential contractors' bidding information $I_{p,b}$ does not include other entities of the environment. Hence, *sharing* $I_{p,b}$ raises privacy concerns. For example, the manger entity realizes about maximum willingness of entities to get a task $I_{p,k}$. This can be exploited for future interactions [56] , [57]. Therefore, the bidding information becomes sensitive in relation with the manager entity.

$$\mathrm{I^S}(\mathrm{I_{p,b}}, e_m)$$

$$\exists\ o_{m,w} \mid\ o_{m,w}(I_{p,b}) \equiv I_{p,k} \wedge \mathrm{I^S}(I_{p,k}, e_m)$$

Where in CNP:

$$\mathrm{S}(\mathrm{I}_{\mathrm{p,b}}, \mathrm{e}_{\mathrm{m}}) \rightarrow \mathrm{D}(I_{p,k}, \mathrm{e}_{\mathrm{m}})$$

When the contractor is awarded with the task, it performs operations on the task information and provides the result of the requested task. He exposure boundary of result description $I_{r,j}$ only includes the contractor entity. In traditional CNP, the result description is submitted to the manager and it forwards it to the requester. This indicates that $I_{r,j}$ is sensitive information in relation with the manager and protection mechanism should be applied.

Entities in CNP can create sub tasks. The default of sub-contractors in CNP is all entities. In another word, all entities will receive the task announcement $I_{r,t}$ and sub task information of the requester. $E_{r,t}$ only includes the manager entity through which $I_{r,t}$ becomes sensitive in relation with subcontractors.

$$\mathrm{I}^{\mathrm{S}}(I_{r,t}, e_s)$$

Because {name, task_announcement} becomes the auxiliary information at the sub contractor entities, it might be used for transforming sensitive implicit information to explicit. Similar to this case is when the $\{name, task\_announcement\}$ is sent to entities that are not capable of executing the task and therefore, they will not be competing over the task. However, this information is used as auxiliary information at $Not\ capable\ contractors: e_n$. Therefore,

$$\mathrm{I}^{\mathrm{S}}(I_{r,t}, e_n)$$

$E_{r,h} = \emptyset$ which indicates the history of tasks $I_{r,h}$ at the manager entity is sensitive. This is due to existing of auxiliary information such as history of allocations of requesters and contractors among entities might be used for retrieving sensitive information. As an example, an entity has requested for temperature controlling services $I_{r,t}$ over a specific area in a higher frequency during the last two days $I_{r,l}$. The manager entity using some auxiliary information I^{aux} such as fire alarm services on that region can transform the implicit information to explicit information that there was a fire accident ($I_{r,k}$) on the

entity's site belonging to that area. This information is sensitive in relation with the manger entity.

$$\exists\ o_{m,w} |\ \ o_{m,w}\left(\{I_{r,t}, I_{r,l}, I^{aux}\}, I_{r,k}\right) \wedge \mathrm{I}^{\mathrm{S}}\left(I_{r,k}, e_m\right)$$

Where in CNP:

$$\mathrm{S}\left(I_{r,t}, \mathrm{e_m}\right) \wedge\ \bar{\bar{o}}_{m,w}\left(\{I_{r,t}, I_{r,l}, I^{aux}\}\right) \rightarrow \mathrm{D}\left(I_{r,k}, \mathrm{e_m}\right)$$

Through this it is deduced that protection mechanisms have to be applied to avoid the disclosure of sensitive information.

Requesters and contractors potentially can be allocated to each other when similar tasks are announced. Because the contractor is capable of storing the received task announcement, it implicitly possesses information about the requester that can be sensitive. The exposure boundary of $I_{r,s}$ does not include any entity $E_{r,s} = \emptyset$. The proposed framework can reason accordingly, that this information is sensitive in relation with the contractor entity. For instance, when the computation analysis of measuring the required resources for the new project is assigned to a contractor using CNP, monitoring their results is a period of time can transform the implicit sensitive information such as stock growth rate of the business in near future. To deal with this, privacy protection mechanism should be applied to avoid operating on history of tasks that are allocated to a contractor.

$$\mathrm{I}^{\mathrm{S}}\left(I_{r,s}, e_c\right)$$

6.3 CNP in privacy protection management framework

In this section, we focus on how the privacy protection management framework transforms the CNP interaction protocol to privacy based CNP (PB_CNP). In the previous subsection, the framework deducted the sensitive information through the exposure boundaries of the information *shared* in traditional CNP. This enables the framework to apply adequate protection mechanism for the identified sensitive information as part of the interaction protocol.

6.3.1 Task Announcement

To avoid *sharing* the combination of task announcement and the identity of the requester $I_{r,t}$, a preventive mechanism at the information level is required. This allows the protocol to *share* the information that is required for executing the task. In addition, it manipulates the information through which the task cannot be attributed to the requester. The sequence of CNP at this level includes the following operations:

CNP:[REQUEST:TASK_ANNOUNCEMENT,CFP:TASKANNOUNCEMENT]

The framework employs the protection mechanism at this sequence before it is *shared* with potential contractors. Among mechanism can be applied at this level are anonymization or cryptographic mechanism. Therefore, it will be transformed to:

PB_CNP:[REQUEST:TASK_ANNOUNCEMENT,ANONYMIZED_TASK,CFP:TASK ANNOUNCEMENT]

However to provide the possibility of de-identifying and anonymization of tasks, there are procedures have to be added to the protocol procedures and message structure should be modified:

- Substituting {$name$} with a task identifier
- Mapping the task identifier and the name of entities only is kept in the manager which is part of the auxiliary information they have about entities
- <task-announcement> ⇒ TASK-ANNOUNCEMENT [Task_id] {task-abstraction} {eligibility-specification} {bid-specification} [expiration-time]

6.3.2 Proposal/Bid

In CNP, the bidding information in conjunction with the identity of the potential contractors will be sent to participate in winner determination in the manager [broker] entity. The combination of the bidding value and the identity of the bidder is sensitive in relation with the manager [56]. This is due to the exposure boundary of this information which is $E_{c,b} = \emptyset$. However, CNP *shares* this information. The CNP in winner determination is as the following:

$$CNP: [Propose, Accept\_Proposal]$$

One of the mechanisms to protect privacy in this context is applying preventive mechanisms at the information level using the mechanism proposed in [57] at which calculation happens on encrypted information and the broker is not aware of the bidding values. Utilizing this mechanism will extend the protocol with additional processes to encrypt the bidding information.

$$PB\_CNP: [\ BidEncryption, Propose\ , Accept\_Proposal]$$

6.3.3 Result Description

$E_{r,j}$ only includes the contractor which deduces that the result information is sensitive in relation with the manager. To protect this information, it is required to prevent the *sharing* operation (preventive mechanism at the operation level). This enforces the contractor to identify the owner of the task and directly send the information to them. Also it is possible to use cryptographic approaches to encrypt the result information with requester public key (preventive mechanism at the information level). In both cases, the contractor has to perform a procedure to realize the owner of the request.

CNP: [RESULT]

PB_CNP: [ONWER_REALIZATION, RESULT]

6.3.4 Subcontractors

The operation of *sharing* and disseminating information to sub contractors may disclose implicit sensitive information. Contractors can send direct messages to other entities to allocate some part of the task to them. Therefore, it is required to perform a preventive mechanism at the operation level. The node abstraction will include the list of sub contractors. The manager can exclude proposals that include subcontractors that do not belong to the exposure boundary of subtask information $I_{r,g}$. For instance, to perform an operation on computational resources, only entities that belong to a particular geographical location are allowed to acquire the task. This inherits to entities that execute the subtasks. In addition, the *sharing* operations can be prevented using approaches such as the work in [72] which only allows execution of operations that are captured as part of

a contract agreed by participants. To apply these mechanisms, it is required that the requester includes the exposure boundary of subtasks $E_{r,g}$ as part of the task announcement.

<task-announcement> ⇒ TASK-ANNOUNCEMENT [Task_id] {task-abstraction} {eligibility-specification} {bid-specification} [expiration-time] [subtask boundary]

Through this a new procedure at the manager will be applied which extends the operations at the protocol.

CNP: [TASK_ANNOUNCEMENT, PROPOSE, ACCEPT_PROPOSAL]

PB_CNP: [ANONYMIZED_TASK, PROPOSE, SUBCONTRACTOR_CHECK, ACCEPT_PROPOSAL]

6.3.5 Non Capable Potential Contractors

Because task information $I_{r,t}$ is sensitive in relation with potentials contractors that are not capable of executing the task, it is required to reduce the task announcement phase to entities that have the capability to execute the task. This requires preventive mechanisms at the operation level. It includes new messages and sequences at the protocol in which potential contractors register their capabilities with the manager. This allows the manager to multicast the task announcement to entities that have the potentials to execute the task.

CNP: []

PB_CNP: [REGISTER,CONFIRM]

This sequence introduces a new procedure at the manager to evaluate the potential contractors before announcing the task.

6.3.6 Task History

In forms on CNP that the manager and requester entity are not the same, the history of task allocation is sensitive in relation with the manager $I^S(I_{r,h}, e_m)$. Because the manager uses auxiliary information to perform operations that may not be authorized, a mechanism at the operation should be performed. The task information is *shared* during a

period of time that conducts to new information when they are used altogether. Through this preventive mechanisms may not be effective and punishing mechanisms is more adequate. The punishing mechanisms obliged to include an agreement where only certain set of operations can be applied. This introduces new information and sequences to extent the protocol. $\{< Agreement, Operations >, < Confirm, Task\ Announcement >, < Refuse >\}$

6.3.7 Result History

Because the contractor can convey the history of the tasks that are allocated to it and perform operations on it, implicit sensitive information might be transformed to explicit using auxiliary information. This information is *shared* with the contractor entity in a period of time. To protect this information punishing mechanisms can be more effective. These mechanisms require having an agreement between the participants of the interaction. This introduced the agreement process between the contractor and the requester (Figure 19).

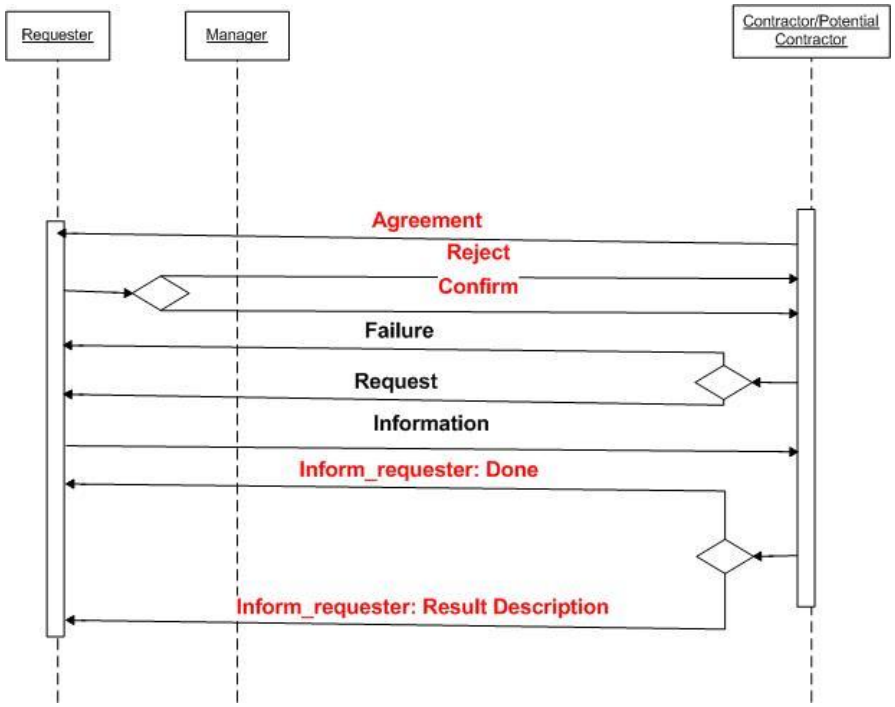

Figure 19. Result history as sensitive information

$\{< Agreement, Operations >, < Confirm, Task\ Announcement >, < Refuse >\}$

6.3.8 REQUIRED information and Information Specification

In traditional CNP, the potential contractors can request for more information from the requester to perform a task. The required information is sent through REQUEST message and its response $I_{r,y}$ is carried with INFORMARTION messages. Depending on the requested information the exposure boundary might change. It could consist the contractor but it also might not contain the contractor entity. Thus, the required information might include sensitive information or it may disclose sensitive information when it is used in combination of task announcement.

$$I^S(I_{r,y}, e_c)$$

The node abstraction is submitted to the manager entity in proposal phase. It includes the possible information that might be requested if the potential contractor becomes the contractor. To deal with this, applying protection mechanism at the operation level is required where the dissemination operation is prevented if the submitted required information does not match with the node abstraction that the requester has release. Through this the requester can incorporate the possible extra information in the structure of node abstraction that potentials entities are part of its exposure boundary into the task information.

<task-announcement> ⇒ TASK-ANNOUNCEMENT [Task_id] {task-abstraction} {eligibility-specification} {bid-specification} [expiration-time] [subtask boundary] [extra_information]

Also, this introduces a new procedure at the manager level to filter the potential contractors that their requested information does not belong to the exposure boundary in extra information.

CNP: [PROPOSE, ACCEP_PROPOSAL]

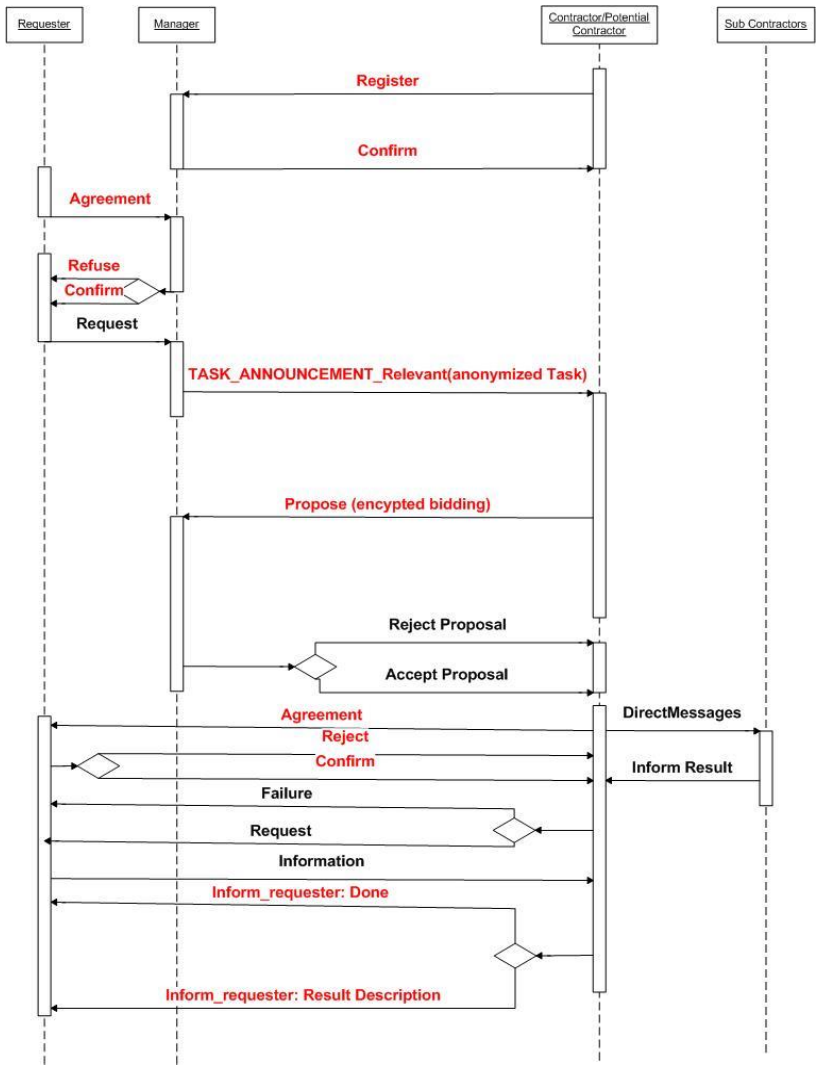

Figure 20. PB_CNP Sequence Diagram

PB_CNP: [PROPOSE, VERIFIED_INFO , ACCEP_PROPOSAL]

And PB_CNP: [PROPOSE, VERIFIED_INFO , REJECT_PROPOSAL]

Given the above modification, the framework generates a privacy based CNP that is depicted in Figure 20.

6.4 Summary

Contract Net Protocol is a negotiation-based protocol used for resolving capability-based interdependency applied in distributed problem solving. Because of the privacy concerns in CNP, it may result in unacceptable solutions. Given the exposure boundary of the information exchanged in the sequences of the protocol, there is sensitive information that is disclosed to other entities. Applying the privacy protection management framework identifies the concern points of the protocol and provides adequate privacy protection mechanisms that creates a privacy-based contract net interaction protocol.

Chapter 7

7 Privacy aware CDS Model: Application Scenarios

Many practical applications can be effectively modeled as CDS environments. They can involve with various services, information sources, devices, equipments and sensors. Internet of Things is one example that can be effectively modeled as CDS. Along this direction Smart-Space is a research initiative at our CDS-Eng research Lab, through which we investigate, several critical research issues, including privacy concerns in open environments. Additionally, two projects within IoT smart space initiative have been included in our research investigation, namely Grid-based resource scheduling and intelligent assistance. In this chapter we elaborate on the feasibility of applying the proposed privacy protection management framework in these application scenarios.

7.1 Smart Space

A smart space project has been implemented as an Internet of Things (IoT) environment in Cooperative Distributed Systems Engineering (CDS-ENG) research lab. It includes sensors, equipment, services and data resources that are exposed to applications. Within this environment, there are entities modeled as agents. Services in this environment utilize the existing resources in the space and deliver solutions to applications. A brokering layer provides functionalities to integrate with resources of the environment including data, services, clouds and events.

7.1.1 Setting of Smart Space

Smart space includes entities with various types of capabilities. They are modeled as agents within the environment. Diverse set of devices, sensors and equipment are used in smart space such as kinects, twines, mindstorm, IP cameras, NFC and RFID tags and android-based mobile devices. The logical architecture of the smart space is shown in Figure 21. Many applications and services are created by utilization of these “things” that are registered within this environment. The “thing” layer mainly refers to the physical devices and application layer encompasses the application and services

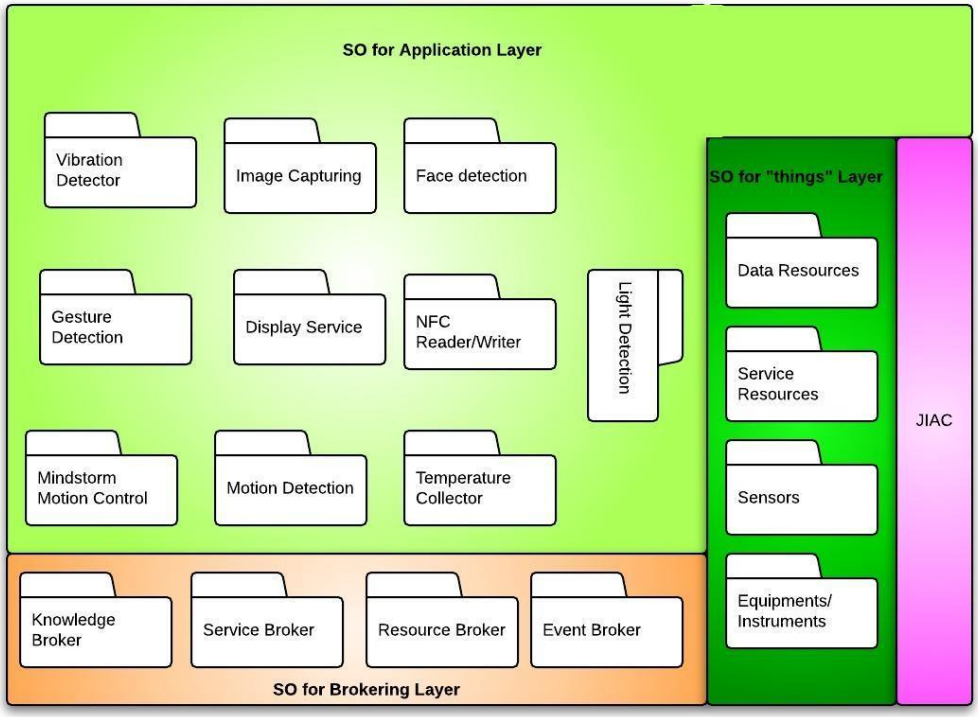

Figure 21. Logical architecture of smart space

presented in smart space. The brokering layer contains four main components to provide knowledge, services, events and resource broker. The Service Broker (SB) is responsible in delivering the requested services to entities. Knowledge Broker (KB) provides a unified view on many data sources and makes it available to entities on their demand. Event Broker (EB) allows entities to register for notifications on occurrence of events to response effectively. The Cloud/Resources Broker (C/RB) provides the scheduling services and resource allocation to entities of the environment. The JIAC platform is providing necessary platform services to accommodate the agents of the environment. In addition, entities in smart space have interdependency problem. To resolve their capability interdependency, we have applied CNP protocol where the brokering entities act as the manager entity.

Entities in smart space are modeled as CIR agents. Because of limited computation capabilities of some of the sensors, the communication part of the CIR-agent model is

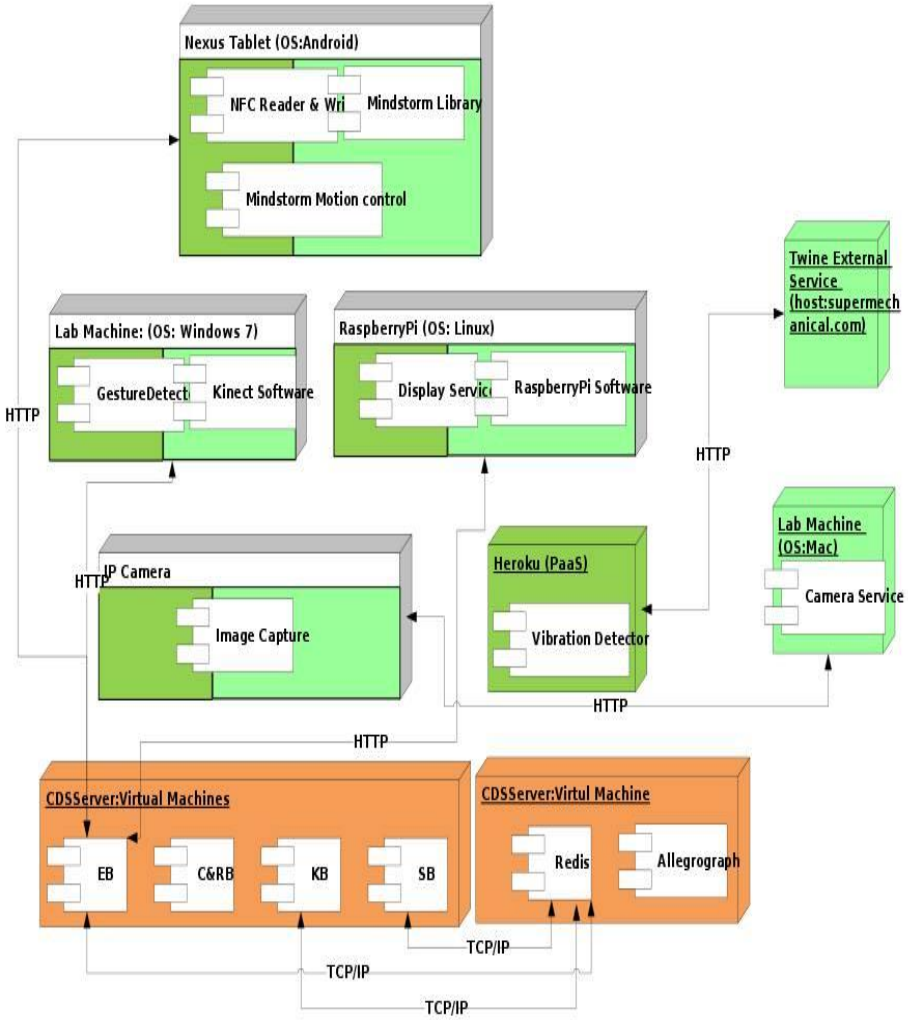

Figure 22. Deployment Diagram

integrated at device and interaction, knowledge and problem solving are incorporated at the components of applications. In result the "thing" layer and application layer are deployed at the same node in some cases. The deployment diagram of smart space is provided in Figure 22.

7.2 Privacy in IoT Environments

The motivation of smart space was to create an IoT environment where various types of "things" could join the environment and utilize and provide services of the environment. The IoT is becoming the newest computation environment with the interconnection of software and information services, devices, equipment, and sensors. These "things" are able to communicate with each other via the Internet [73]. The future growth of IoT based applications is foreseen to be tremendous [74]. The incorporation of social networks and ubiquitous computing technologies in IoT enables individuals and groups of people to interact seamlessly with the environment [75], [76], [77]. The comfort experienced via innovative technologies in IoT is with the expenses of privacy [2], [3], [78]. The more one engages with IoT based applications and their enabling technologies, the more privacy concerns arise [79], [80], [81]. As an example, magnet sensors enable opening doors through the internet. However, for security reasons, they are connected to video sensors which authorize people at the entrance. Applying facial recognition programs on videos combined with the frequency of appearances of people at the house front, may identify members of the family, including children. Using Facebook's facial recognition software also makes it possible to find their Facebook profiles and, possibly, the school that they are going to [82], [83].

As the smart space inherits the characteristics of IoT, privacy becomes a challenge within this environment. Since IoT is modeled as CDS, we have applied the privacy protection management framework at the interaction protocols applied in this space. However, the "things" might refer to small sensors that do not have the computation power to manage the requirements of the privacy-based interaction protocol. Due to limited capacity in some "things" of the environment, we have deployed the interaction capabilities on the nodes in which the components coexist with the components of application layer as depicted in Figure 22. In addition, the traditional CNP was replaced with the privacy-based CNP presented in Chapter 6.

7.3 Privacy based Scheduling Protocol in Smart Space

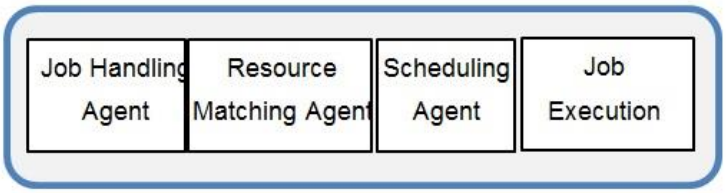

Figure 23. Resource Broker High Level View

The resource broker in smart space provides scheduling and resource allocation services to the entities of the environment. The resources can include grids, clouds and single entities with computation power through which various cloud providers including IaaS, PaaS and SaaS can be connected to the smart space. The cloud-computing paradigm can be classified as IaaS (Infrastructure as a Service), PaaS (Platform as a Service), or SaaS (Software as a Service), where the resources in terms of infrastructure, platform, and software are provisioned as services [84], [27], [85]. Cloud environments can encompass various types of entities from a grid entity to smaller sensors [86]. Figure 23 shows the high level view of the architecture of the resource broker in smart space. The scheduling mechanism in resource brokering is extracted using the work in [87] and the scheduling interaction is presented in Figure 24.

Due to exchange of information in interactions with the resource broker in smart space, privacy becomes a concern at the scheduling interactions. Consider the following example: in a CDS, we have two entities, e_i and e_j, that provide computational resources. e_k and e_y are two entities that require computational resources. e_b is a resource scheduler that allocates resources to entities based on their time and price boundaries. If e_b allocates e_i as the resource provider for e_k, then e_j becomes the only available resource provider. If the job that e_k has *shared* with e_b reaches to e_j, then e_j knows it does not have competitors and it increases the price of the service. This affects the job that e_y has posted to e_b. Either the price is not in the range of e_y price boundaries and the job cannot be done, or it has to be delayed until e_ks job execution is finalized. This scenario one of

the forms of privacy concerns that is referred as “Price Discrimination”[88] , [89]. This is due to the unauthorized disclosure of information to e_j. In this case, e_b caused privacy concerns for e_y as well as e_i. Given such an environment, it is essential that entities receive privacy protection when coordinating with each other in resolving schedules.

Within the scheduling interaction protocol in resource broker, combinations of tasks and the identities of entities can disclose information about those entities. Operating on tasks that are attributed to an entity can reveal information regarding the pattern of the work in the entity, as well as identify highly loaded time slots of a certain entity. This information can be used to implement more effective DoS or DDoS attacks against the entity. This indicates the need for a preventive protection mechanism at the information level, such as anonymization.

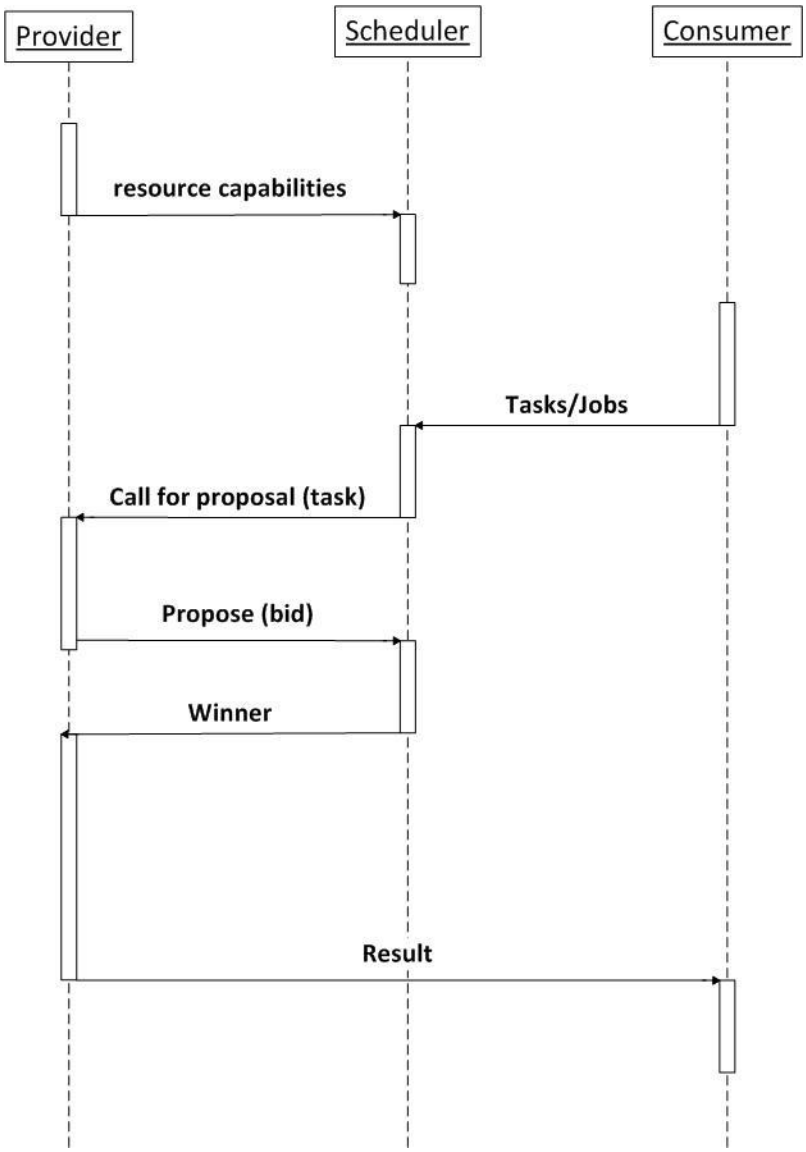

Figure 24. . Scheduling interaction protocol in resource broker in smart space

Similarly, a combination of "propose message" and the bidding value may become sensitive when it discloses the maximum willingness of an entity to acquire a task. For example, in a second price auction mechanism, the auctioneer can start an auction with higher value when they are aware of the existence of an entity that is willing to pay the proposed value. One mechanism that protects the information of proposed messages is the application of the preventive protection mechanism introduced in [57].

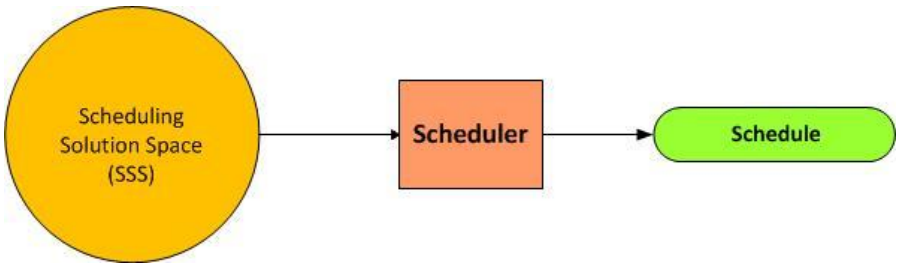

Figure 25. Scheduling solution space

7.3.1 Privacy-based Scheduling Solution

The approach in resolving the privacy problem as a quality factor in scheduling is intended to limit the solution space for entities that can provide the necessary privacy protection. As shown in Figure 25, the scheduling solution considers all entities as part of the solution space. Then it identifies the entity that can resolve a scheduling request.

PPL and the risk measures are the parameters that can be used as decision variables of scheduling solution or as constraint variables in the solution space. In the presented work, we focused on the latter. Figure 26 illustrates that the privacy protection mechanism reduces the solution space by eliminating the interactions that do not have the requested

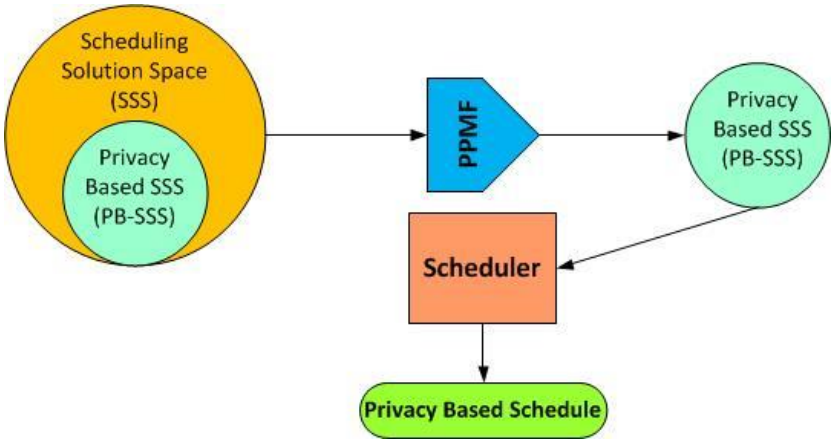

Figure 26. Privacy based scheduling solution space

PPL and risk value.

By applying protection mechanisms of anonymization [64] and private-bid-communication [57], the privacy based scheduling interaction is achieved as depicted in Figure 27. These operations can be substituted with other protection mechanisms such as M_1 and M_2, with protection levels PPL_1 and PPL_2, that serve the desired level of PPL. Applying these mechanisms in the privacy protection management framework results in a privacy based interaction protocol with protection level $PPL_1 * PPL_2$. The risk value is calculated once the job is submitted to the scheduler. This encourages the scheduler not to consider entities with high levels of risk in disclosing the information.

As the privacy-based scheduling interaction protocol reduces the solution space to entities that privacy is protected with a certain degree, the scheduling mechanism can rely on the assumption that privacy is protected. Therefore, any solution that the scheduler provides is within the privacy-based solution space and thus it is acceptable.

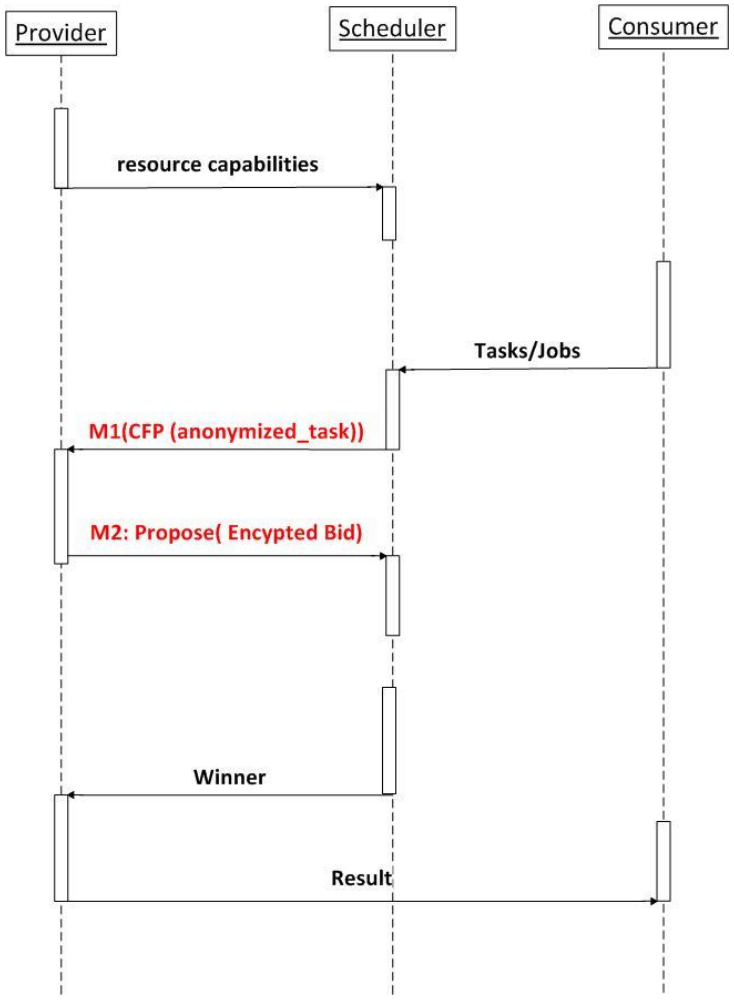

Figure 27. Privacy-based Scheduling Interaction Protocol

7.4 Privacy in Personal Assistant

The explosion of online, mobile, and social networks transformed computation into a platform that redefines many aspects of our personal and business lives. As a result, many of our goals are technology driven and, for us to be able to achieve these goals, we might need to go through several applications and services. Within smart space we provide intelligence assistant that follows the models of personal assistant for the users of the environment. The architecture of the smart assistant in smart space within the CIR agent is depicted in Figure 28 [90]. Interaction is managed in the environment model, and the proposed privacy framework is applied in interaction protocols at this level.

Personal assistants are smart software agents that provide users with services that can adapt to the user and the environment [91]. A cognitive user model is an essential component in personal assistants. Our focus is on proposing a user behavior model that captures users' behaviors in open environments [92], [93].

In this context, users' privacy is a concern when interacting with other entities in smart space through their PAs (Personal Assistants). Users' interests are considered sensitive information, which might be disclosed to other entities [59], [60]. For instance, promoting movies, books, software, web sites, and other products regarding a particular heritage, religion, or group of people with a *shared* political or social opinion, demonstrates users' interests about these topics, which may be a privacy concern. Identity and interest information about users can be used as implicit information in conjunction with other operations.

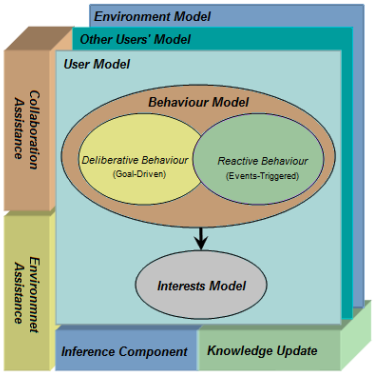

Figure 28. Personal Assistant Architecture

7.5 Privacy Based Personal Assistant

By *sharing* explicit information of users' interests (such as grocery items, favourite books, etc), we might implicitly disclose sensitive information. Providers may exist in the boundary of users' interest information. However, the exposure boundary of users' interest information, such as their favourite book, does not include entities offering jobs or insurance. Users' interest in horror story book may prevent them from getting job opportunities in nursing. Therefore, employers do not belong to the exposure boundary of

favourite book information. Similarly, purchasing cancer treatment books may imply that the user may suffer from cancer. This information is sensitive in relation with insurance recommendations. In Smart space, entities apply anonymization and pseudonymity with adequate levels of PPL when interacting with entities that provide assistance. Applying the privacy protection management framework within the personal assistant reduces the assistances that are offered to the user for which the privacy is considered. However, depending on the exposure boundary of interest information, the assistance is filtered by the privacy-based interaction protocol.

7.6 Summary

In CDS-Eng research lab, a smart space project is defined in which includes diverse set of sensors, equipment and devices to form an IoT environment and is modeled as CDS. “things” in smart space are modeled as agents that interact with other agents in smart space. Because of the increased involvement of people and their devices in IoT applications, privacy has become a more complex challenge. Hence, we have applied privacy-based interaction protocol in smart space.

One component of the brokering layer in smart space is resource broker that delivers scheduling service using scheduling interaction protocol. However, privacy within this setting is still a concern. As entities in these environments are autonomous and self-interested, it is assumed that all entities will respect privacy. Any scheduling solution that results in privacy concerns will not be acceptable. Therefore, privacy becomes a quality factor of the scheduling solution. The interaction of entities in scheduling will be transformed through privacy-based scheduling interaction protocols which enable scheduling engines to delegate privacy concerns to interaction protocols. The scheduler can assume that all entities respect privacy, as the interaction protocol has reduced the solution space to privacy protected boundaries.

Intelligent assistants in smart space are provided through personal assistants that interact with the environment to provide relevant assistance. Due to sensitivity of interest information of users, personal assistants of smart space apply the privacy protection

management through which it enables users to receive relevant assistance from providers that their interaction conveys acceptable level of privacy protection.

Chapter 8

8 Conclusion and Future Work

The goal of this research is to provide a formal treatment of "privacy" as a fundamental computation concept in CDS paradigm to build a privacy-aware CDS framework and platform. The formal model of privacy served as a base for developing a privacy protection management framework for CDS. It includes a privacy-aware agent model for CDS platform with the ability to support interaction-based privacy protection. Additionally, the feasibility of the proposed models has been demonstrated by developing an agent-based CDS platform using JIAC framework in an IoT-based project of smart space and a privacy-based Contract Net Protocol.

8.1 Summary of contributions

An important class of distributed systems is CDS, in which entities are able to exercise some degree of authority in *sharing* their capabilities. Entities in this paradigm are expected to cooperate to achieve individual or collective goals. Due to interdependency problem among entities, they require the coordination of their activities using interactions. In the message-based form of interactions, entities exchange information through autonomous and self-interested entities, and thus their privacy becomes a concern. In CDS, solutions are accomplished through the participation of several entities where each has only part of the solution. This positions CDS as a computation platform in which the computation occurs at entities' interactions. This entails that privacy challenges in CDS are the concerns associated to the computation happening at the interaction level.

8.1.1 Challenges and Contributions

Privacy, by nature is a concept that is defined with many denotations, which could be interpreted differently in various contexts. Understanding privacy as a concept that can be applied in contexts such as CDS requires formal analysis of settings in which privacy is not negligible. Despite existing privacy models proposed in many contexts, attendance

to privacy models that capture privacy at computation and is adequate for CDS is lacking. To address privacy as part of the quality of the solutions that are reached in CDS, a certain degree of the privacy protection should be guaranteed during entities' interaction. Furthermore, an approach is required to properly apply the privacy model at the interaction level to enable privacy protection as part of the computation in CDS.

The perspective about privacy in related works can be categorized into two major areas. First is verifying the legitimacy of the achieved solution after applying the privacy constraining rules. Secondly, incorporating privacy in the solution as a computation concept. To resolve privacy concerns in CDS, it is essential that privacy is modelled in a context that is adequate for CDS environments. There are many related works that have addressed privacy in contexts that are not capable of encompassing the complete settings of applications in CDS environments. For instance, the differential privacy and its affiliated applied mechanisms are aimed for statistical data analysis contexts. CDS is a broader area where it is essential to provide a privacy model that is applicable in it. Modeling privacy in information management context can be categorized as information collection, information processing and information dissemination through which it can adequately be applied in CDS environments. Modeling the solution for privacy is typically classified as rule-based and architectural-based approaches. Due to inconveniences of rule-based approaches in dynamic environments, architectural-based approaches are more desirable for CDS environments. In this work, we pursue the computation view on privacy within the information management context and adopt the architectural-based solution approaches by applying the model at the interaction level.

8.1.2 Formal Modeling of privacy

Due to lack of formal analysis on privacy that is adequate for CDS, in this work we proposed a formal model for privacy in an information management context. Privacy is the concern of decentralized environments where the control and knowledge are distributed among autonomous, self-interested entities and they need to adopt message-based interactions through which information is *shared*. *Sharing* is a supervised process by entities, and as such depending on the receiver of the information, the entity does not *share* the information that is classified as *sensitive*. Information can be sensitive in

relation to an entity and become non-sensitive in relation to another. It creates an exposure boundary for any information that the entity possesses. This entails privacy is the state of the exposure boundary for which information is not sensitive when it flows within the exposure boundary and it becomes sensitive when it is outside this boundary. Information exists as explicit forms, however, it can be implicitly available when information is used in conjunction with operations. Manipulation of information by operations can transfer the information outside of their exposure boundary. The "disclosure" of information refers to explicitly or implicitly making the information available at the receiver entity. The "privacy concern" within this context relates to disclosing sensitive implicit information. Any operation that transforms the sensitive implicit information to explicit form becomes non-authorized. The Execution of non-authorized operations when there is agreement between the two parties to not apply the operation refers to privacy violation. Preventing or neutralizing the execution of non-authorized operations becomes the privacy protection definition.

The security mechanisms are mainly applied at the exposure boundary controlling the sharing process. However, these mechanisms are not sufficient to manage the disclosure of sensitive implicit information, which happens outside of the exposure boundary.

Perfect privacy protection happens when all non-authorized operations are prevented or neutralized. Due to the incomplete knowledge of entities in CDS, perfect protection might not be attainable and quasi protection mechanisms will be applied. To address the uncertainty level of privacy protection in quasi mechanisms, a probabilistic model is proposed that reflects conditional probability of privacy protection given the information that exists at the entity. This concept is addressed as Privacy Protection Level (PPL).

The trust concept is the degree of entities' belief over the reliability of an entity on executing or not executing certain set of operations. Therefore, this concept can impact the level of probability that the unauthorized operations are applied which is reflected on PPL value.

8.1.3 Solution for privacy within CDS environments

Many computation approaches are applied in some settings of CDS environments for which not all CDS applications can adopt them. Providing the privacy concern solution at the interaction level allows for the adoption of the framework by CDS-based environments. The proposed privacy protection framework models the protection mechanisms as operations and classifies them at preventive and punishing mechanisms. The protection mechanisms can be applied to the information by distorting or altering the information. It also can happen to the operation by not allowing certain operations to be executed. Interactions between entities can be captured in information management and be categorized as information collection, processing and dissemination. The interaction protocol can be modeled as a set of messages and a set of sequences of exchange of messages. Furthermore, interaction protocols can be modeled as a set of operations that are executed in the sequences presented in the interaction protocol. By applying the privacy model at the interaction protocol, the adequate privacy protection mechanism is added to the operations of the interaction protocol. In the end, the framework expands the interaction protocol with proper messages and sequences that reflects on the protection that is applied on the interaction protocol.

It is proven in this work that protection at the interaction protocol is sufficient for protecting privacy in CDS environments. Also, the generated privacy-based interaction protocol has quantifiable privacy protection level that allows entities to interact with a certain degree of protection.

Entities in CDS have an interdependency problem for which they need to interact in order to reach to a solution. The computation entity within CDS can be adequately be modeled as CIR agents that have knowledge, problem solving capabilities, interactions and communications. The solution within such a computation platform is achieved by conjunction of problem solving and coordination solution, which can be managed by interactions. However, the solution that is achieved might not be acceptable if the privacy concerns are not resolved. Within this work, we have applied the privacy protection management framework at the computation level by expanding the structure of the entity to include privacy protection management that adheres to the privacy-based interaction

protocol. In this work, we have formally explained that legitimate acceptable solutions at the computation require the inclusion of privacy resolution in-addition to problem solving and coordination.

8.1.4 Privacy-Based Contract net protocol

Contract Net Protocol (CNP) is a negotiation-based interaction protocol that is used for distributed problem solving. Entities within CNP can be autonomous and self-interested, allowing this protocol to be associable in CDS. Application of the privacy models and the proposed privacy protection management framework on CDS identifies the privacy concerns related to this protocol. By utilizing the proposed framework as an analytical tool as well as applying it at the computation entity, it is possible to expand this protocol with the necessary privacy protection operations.

8.1.5 Implementation Challenges

To implement the privacy-aware computation, we have used the JIAC agent platform that provides the necessary functionalities for communication and agent life cycle management. This platform does not include the interaction mechanisms at the agent level and everything is tailored to actions within agents. To resolve interactions at the JIAC platform, we have introduced the interaction as an agent bean that follows the sequences and messages it receives. Sequence management is performed by providing a state management class that creates and monitors the states of the interaction protocol. The challenge regarding the expansion of the interaction protocol with protection operations is resolved by introducing protection operations that are registered with the knowledge of the entity and at the same time are introduced as available actions. To enable adaptation of the interaction protocol with various PPL levels and different interaction protocol, the actions are added to the agent node memory before the entity gets to the start state at the initialing phase. The protection operations that are simulated within the JIAC platform are pseudonymity, private bid transfer, early registration and applying agreements. The PINQ platform [95] also is one of the anonymizer operations that can be applied as protection operations.

The proposed privacy framework can be applied in many applications that are modeled by CDS. In this work, we have shown the project of smart space that is an IoT-based environment that is modeled as CDS and the proposed privacy protection management framework is applied to them. Similarly, the scheduling interaction protocol within the resource broker of smart space is substituted with privacy-based scheduling interaction protocol that reduces the scheduling solution space to the ones where privacy is respected. In another example, the proposed privacy protection framework is applied in personal assistance applications where users' interests are sensitive information and still are *shared* with the assistant providers.

Although a level of privacy protection is achievable within the presented work, privacy violation is bound to disobeying the agreement among participant entities. The agreement includes operations that are considered to be non-authorized. Theoretically, within an environment, any state changes if an operation is applied. This indicates that all the operations can be addressed by capturing the states of the environment. However, from a practical perspective, it is envisioned to have environments in which knowing the non-authorized operations might not be possible for the entities. Therefore, the entity would not be able to set proper agreements. In any privacy based interaction, there is $(1 - PPL)$ chance that the privacy protection is not provided which entails the probability of transforming the implicit information to explicit by execution of non-authorized operations. Evidently, not having adequate mechanisms to avoid privacy violation impacts the level of risks within interaction.

8.2 Future Work

Our contributions were mainly in the areas of modeling and categorizing privacy, formal analysis of privacy within information management, computational view on privacy at the interaction protocol and providing privacy aware computation systems. However, this approach can be expanded within the areas of economic-based privacy model and optimization of privacy protection management.

- **Privacy Protection Management in Computation**

Due to the incomplete knowledge of entities in CDS, attaining perfect protection with adhering to preventive mechanisms is challenging. This drives the application of quasi privacy protection mechanisms that are subject to levels of uncertainty on strength or the confidence factor of the applied mechanisms. The uncertainty associated with protection mechanisms are captured by measuring the PPL value. The probability inherited in this merit reflects the probability of the protection and the probability of privacy concerns occurring. As an example if PPL=0.75, there will be 25% chance that the privacy concerns will occur.

The decision-making process happening at the entity level can accept the chance of a privacy concern by measuring the risk of interaction. This is in direct relation with utility and the accepted level of PPL in general while interacting with other entities. Therefore, it becomes a multi-objective problem to allocate proper protection operations with an adequate level of PPL which is serving the expected utility and requested protection. The remaining questions concern understating the relationship between the requested PPL and the utility that is expected as well as the risk elements that might impact the requested PPL.

- **Risk analysis**

The risk of interactions related to the probability of occurrence of the negative impact of the privacy concerns entities. It has been captured as a probabilistic model which can include several parameters involved. It is clear that the risk of interaction has significant impact on decision-making process of entities. Computationally, this concept has to be captured at the interaction level. Identifying and encapsulating the parameters impacting the risk of interaction and extend the interaction protocols to adhere to risk analysis in one of the questions that we are going to answer in future works of this research. Furthermore, optimizing the level of risks, the level of protection and the level of utility to serve entities objective in their interactions still requires investigation.

- **Punishing mechanism**

The punishing classes of protection mechanisms are technically mitigation operations that reduce the desire to violate privacy at the receiver side due to the consequences that might be imposed on them by their agreements. The effect of the incentives or the punishments that are integrated as part of the agreement can be measured by an analysis of the probabilities of existing strategies for the entities. This becomes an approach for evaluating the PPL associated with punishing mechanisms. One of the approaches towards the agreements of the punishing mechanisms is through the economic-based modeling of the privacy-based interactions through which the punishing mechanisms provide sufficient incentives for complying with the agreement. Nevertheless, punishing mechanisms in the context of an economic mechanism requires further analysis and research challenges that need to be addressed.

8.2.1 Areas of Expansion

- **Economic-based Privacy Model**

Economic mechanisms are adequate models for managing interactions in decentralized systems. There have been several attempts to apply economic mechanisms to solve complex decision problems in CDS [87], [96]. In this work, privacy is captured at the interaction level where decisions to resolve interdependency problems are made. Also, it is observed that entities decide on the exposure boundary based on their evaluation of the utility and possible privacy impact [14, 66]. Applications of quasi protection mechanisms convey certain levels of probability that the privacy concerns may occur after *sharing* information. In economic mechanisms, the dominant strategies are the mechanisms that the best possible choice of entities is what the mechanism is aimed for. For instance, in second price auction, the mechanisms provide incentive for entities to reveal their true valuation [97]. Modeling privacy using economic based approaches can provide alternatives in which entities willingly consider the privacy of others. Because entities are economically rational, the expected outcome is the elimination of the chance of executing operation that transforms non-sensitive information into sensitive. Therefore, the solution to privacy can behave as perfect protection mechanisms.

- **Semantic-Driven Privacy Protection Management**

Entities in open environments require the *sharing* of a similar understanding of the reality and the concepts defined in their knowledge to interact. Due to differences on the semantics that entities adopt, semantic integration is essential quality factor for open environments. This also includes agreeing on the semantics of the interaction protocol. Therefore, the privacy protection management framework requires resolving the semantic integration at the interaction protocol before performing the protection operations. This opens a new perspective on the proposed model that is considered in the future works of this research.

References

[1] D. Lake, R. Milito, M. Morrow and R. Vargheese. Internet of Things: Architectural Framework for eHealth Security. 2013.

[2] S. Poslad, M. Hamdi and H. Abie. Adaptive security and privacy management for the internet of things (ASPI 2013). Presented at Proceedings of the 2013 ACM Conference on Pervasive and Ubiquitous Computing Adjunct Publication. 2013, Available: http://doi.acm.org/10.1145/2494091.2499770. DOI: 10.1145/2494091.2499770.

[3] A. Ukil, S. Bandyopadhyay, J. Joseph, V. Banahatti and S. Lodha. Negotiation-based privacy preservation scheme in internet of things platform. Presented at Proceedings of the First International Conference on Security of Internet of Things. 2012, Available: http://doi.acm.org/10.1145/2490428.2490439. DOI: 10.1145/2490428.2490439.

[4] (June 2012). HOW MUCH DATA IS CREATED EVERY MINUTE?. Available: http://www.visualnews.com/2012/06/19/how-much-data-created-every-minute/.

[5] A. Krause and E. Horvitz. A utility-theoretic approach to privacy and personalization. Presented at Proceedings of the 23rd National Conference on Artificial Intelligence - Volume 2. 2008, Available: http://dl.acm.org/citation.cfm?id=1620163.1620256.

[6] A. Huertas Celdran, F. J. Garcia Clemente, M. Gil Perez and G. Martinez Perez. SeCoMan: A semantic-aware policy framework for developing privacy-preserving and context-aware smart applications. *Systems Journal, IEEE PP(99)*, pp. 1-14. 2014. . DOI: 10.1109/JSYST.2013.2297707.

[7] I. Kayes and A. Iamnitchi. Aegis: A semantic implementation of privacy as contextual integrity in social ecosystems. Presented at Privacy, Security and Trust (PST), 2013 Eleventh Annual International Conference On. 2013, . DOI: 10.1109/PST.2013.6596041.

[8] Paul M. Schwartz, Daniel J. Solove. The PII problem: Privacy and a new concept of personally identifiable information. 2011.

[9] B. Lang , I. Foster , F. Siebenlist , R. Ananthakrishnan , T. Freeman. A Multipolicy Authorization Framework for Grid Security. Proceedings of the Fifth IEEE Symposium on Network Computing and Application, 2006.

[10] S. Spiekermann and L. F. Cranor. Engineering privacy. *IEEE Trans.* Software Eng. 35(1), pp. 67-82. 2009 . DOI: http://doi.ieeecomputersociety.org/10.1109/TSE.2008.88.

[11] B. K. Sy, A. Ramirez and A. P. K. Krishnan. Secure information processing with privacy assurance - standard based design and development for biometric applications. Presented at Privacy Security and Trust (PST), 2010 Eighth Annual International Conference. 2010, . DOI: 10.1109/PST.2010.5593255.

[12] A. Machanavajjhala, D. Kifer, J. Gehrke and M. Venkitasubramaniam. L-diversity: Privacy beyond k-anonymity. *ACM Trans.Knowl.Discov.Data 1(1),* 2007. Available: http://doi.acm.org/10.1145/1217299.1217302. DOI: 10.1145/1217299.1217302.

[13] N. Li, T. Li and S. Venkatasubramanian. T-closeness: Privacy beyond k-anonymity and l-diversity. Presented at Data Engineering, 2007. ICDE 2007. IEEE 23rd International Conference. 2007, . DOI: 10.1109/ICDE.2007.367856.

[14] T. Li and N. Li. On the tradeoff between privacy and utility in data publishing. Presented at KDD '09: Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. 2009 . DOI: http://doi.acm.org.library.capella.edu/10.1145/1557019.1557079.

[15] R. Dong, A. A. Cardenas, L. J. Ratliff, H. Ohlsson and S. S. Sastry. Quantifying the utility-privacy tradeoff in the smart grid. *CoRR abs/1406.2568*2014. Available: http://arxiv.org/abs/1406.2568.

[16] J. M. Such, A. Espinosa and A. García-Fornes. A survey of privacy in multi-agent systems. Knowl. Eng. Rev. 2012.

[17] C. Clifton and T. Tassa. On syntactic anonymity and differential privacy. 2013 IEEE 29th International Conference on Data Engineering Workshops (ICDEW) 0pp. 88-93. 2013. . DOI: http://doi.ieeecomputersociety.org/10.1109/ICDEW.2013.6547433.

[18] A. Singla, E. Horvitz, E. Kamar and R. W. White. Stochastic privacy. Presented at Proc. Conference on Artificial Intelligence (AAAI). 2014, .

[19] H. H. Ghenniwa. Coordination in Cooperative Distributed Systems, 1996.

[20] M. Tierney and L. Subramanian. Realizing privacy by definition in social networks. Presented at Proceedings of 5th Asia-Pacific Workshop on Systems. 2014, Available: http://doi.acm.org/10.1145/2637166.2637232. DOI: 10.1145/2637166.2637232.

[21] C. N. and S. J.M. Implicit contextual integrity in online social networks. *ArXiv E-Prints* 2015. Available: http://adsabs.harvard.edu/abs/2015arXiv150202493C.

[22] A. Datta, J. Blocki, N. Christin, H. DeYoung, D. Garg, L. Jia, D. Kaynar and A. Sinha. Understanding and protecting privacy: Formal semantics and principled audit mechanisms. Presented at Proceedings of the 7th International Conference on Information Systems Security. 2011, Available: http://dx.doi.org/10.1007/978-3-642-25560-1_1. DOI: 10.1007/978-3-642-25560-1_1.

[23] C. Dwork. Differential privacy: A survey of results. Presented at Proceedings of the 5th International Conference on Theory and Applications of Models of Computation. 2008, Available: http://dl.acm.org/citation.cfm?id=1791834.1791836.

[24] M. Gruteser, J. Bredin and D. Grunwald. Path privacy in location-aware computing. 2004.

[25] D. J. Solove, M. Rotenberg and P. M. Schwartz. Privacy, Information and Technology 2006Available: http://books.google.ca/books?id=Ze3_NDCHK2IC.

[26] D. J. Solove. Nothing to Hide: The False Tradeoff between Privacy and Security 2011Available: http://books.google.ca/books?id=UUdQi4FxRxAC.

[27] B. Kepes, "Understanding-the-Cloud-Computing-Stack," *Rackspace White Paper,* 2011.

[28] JIAC Development Team, "Manual JIAC : Java Intelligent Agent Componentware ," 2014.

[29] D. J. Solove. Understanding Privacy 2008(v. 10). Available: http://books.google.ca/books?id=XU5-AAAAMAAJ.

[30] M. Tentori, J. Favela and M. D. Rodriguez. Privacy-aware autonomous agents for pervasive healthcare. IEEE Intelligent Systems 21(6), pp. 55-62. 2006. Available: http://dx.doi.org/10.1109/MIS.2006.118. DOI: 10.1109/MIS.2006.118.

[31] H. Vanchinathan, G. Bartok and A. Krause. Efficient partial monitoring with prior information. Presented at Neural Information Processing Systems (NIPS). 2014, .

[32] K. Nissim, R. Smorodinsky and M. Tennenholtz. Approximately optimal mechanism design via differential privacy. Presented at Proceedings of the 3rd Innovations in Theoretical Computer Science Conference. 2012, Available: http://doi.acm.org/10.1145/2090236.2090254. DOI: 10.1145/2090236.2090254.

[33] Y. D. Wang. Ontology-driven semantic transformation for cooperative information systems. 2009.

[34] E. L. Godkin, *Libel and its Legal Remedy.* pp. 80, 1880.

[35] S. D. Warren and L. D. Brandies, *The Right to Privacy.* 1890.

[36] R. A. Posner, *The Economics of Justice.* 1981.

[37] A. Westin, *Privacy and Freedom.* 1967.

[38] S. Bok, *Secrets: On the Ethics of Concealment and Revelation.* 1983.

[39] M. Bezzi. An information theoretic approach for privacy metrics. *Trans.Data Privacy 3(3)*, pp. 199-215. 2010. Available: http://dl.acm.org/citation.cfm?id=2019307.2019309.

[40] P. M. Schwartz and D. J. Solove. The PII Problem: Privacy and a New Concept of Personally Identifiable Information. 2011.

[41] S. Lederer, A. K. Dey and J. Mankoff. A conceptual model and a metaphor of everyday privacy in ubiquitous. University of California at Berkeley. Berkeley, CA, USA. 2002.

[42] S. Singh and S. Bawa. A privacy, trust and policy based authorization framework for services in distributed environments.

[43] L. Sweeney. K-anonymity: A model for protecting privacy. *Int.J.Uncertain.Fuzziness Knowl.-Based Syst. 10(5)*, pp. 557-570. 2002. Available: http://dx.doi.org/10.1142/S0218488502001648. DOI: 10.1142/S0218488502001648.

[44] H. Tian and W. Zhang. Privacy-preserving data publishing based on utility specification. Presented at Social Computing (SocialCom), 2013 International Conference On. 2013, . DOI: 10.1109/SocialCom.2013.24.

[45] J. Byun, T. Li, E. Bertino, N. Li and Y. Sohn. Privacy-preserving incremental data dissemination. *J.Comput.Secur. 17(1)*, pp. 43-68. 2009. Available: http://dl.acm.org/citation.cfm?id=1517343.1517345.

[46] M. Yokoo, E. H. Durfee, T. Ishida and K. Kuwabara. The distributed constraint satisfaction problem: Formalization and algorithms. *Knowledge and Data Engineering, IEEE Transactions On 10(5)*, pp. 673-685. 1998. . DOI: 10.1109/69.729707.

[47] B. Faltings, T. Leaute and A. Petcu. Privacy guarantees through distributed constraint satisfaction. Presented at Web Intelligence and Intelligent Agent Technology, 2008. WI-IAT '08. IEEE/WIC/ACM International Conference On. 2008, . DOI: 10.1109/WIIAT.2008.177.

[48] A. Petcu and B. Faltings. DPOP: A scalable method for multiagent constraint optimization. Presented at IJCAI 05. 2005, .

[49] R. Greenstadt, J. P. Pearce and M. Tambe. Analysis of privacy loss in distributed constraint optimization. Presented at Proceedings of the 21st National Conference on Artificial Intelligence - Volume 1. 2006, Available: http://dl.acm.org/citation.cfm?id=1597538.1597642.

[50] R. Greenstadt. An analysis of privacy loss in k-optimal algorithms. Presented at In DCR. 2008, .

[51] G. Weiss, Multiagent Systems: A Modern Approach to Distributed Artificial Intelligence 1999.

[52] H. Lee and M. Stamp. An agent-based privacy-enhancing model. 2008, .

[53] L. Cranor, M. Langheinrich, M. Marchiori, M. Presler-Marshall and J. Reagle. The Platform for Privacy Preferences. 1.0 (P3P1.0) Specification 2002.

[54] L. Crepin, Y. Demazeau, O. Boissier and F. Jacquenet. Sensitive data transaction in hippocratic multi-agent systems. *5485*pp. 85-101. 2009. Available: http://dx.doi.org/10.1007/978-3-642-02562-4_5. DOI: 10.1007/978-3-642-02562-4_5.

[55] J. Niu and S. Parsons. An investigation report on auction mechanism design. CoRR abs/0904.12582009.

[56] M. Naor, B. Pinkas and R. Sumner. Privacy preserving auctions and mechanism design. Presented at Proceedings of the 1st ACM Conference on Electronic Commerce. 1999, Available: http://doi.acm.org/10.1145/336992.337028. DOI: 10.1145/336992.337028.

[57] I. Damgard, M. Geisler and M. Kroigard. Homomorphic encryption and secure comparison. *Int.J.Appl.Cryptol. 1(1)*, pp. 22-31. 2008. Available: http://dx.doi.org/10.1504/IJACT.2008.017048. DOI: 10.1504/IJACT.2008.017048 .

[58] A. Samani, H. H. Ghenniwa and J. Samarabandu. Risk-based modelling for managing privacy protection. Presented at Electrical & Computer Engineering (CCECE), 2012 25th IEEE Canadian Conference On. 2012, .

[59] V. Toubiana, H. Nissenbaum, A. Narayanan, S. Barocas and D. Boneh. Adnostic: Privacy preserving targeted advertising , 2010.

[60] A. Juels. Targeted advertising ... and privacy too. Presented at Proceedings of the 2001 Conference on Topics in Cryptology: The Cryptographer's Track at RSA. 2001, Available: http://dl.acm.org/citation.cfm?id=646139.680791.

[61] N. Nagaratnam , J. Dayka , A. Nadalin , F. Siebenlist , V. Welch , I. Foster , S. Tuecke, "The Security Architecture for Open Grid Services," 2002.

[62] C. Dwork, F. McSherry, K. Nissim and A. Smith. Calibrating noise to sensitivity in private data analysis. Presented at Theory of Cryptography, Third Theory of Cryptography Conference, TCC 2006. 2006, Available: http://www.iacr.org/cryptodb/archive/2006/TCC/3650/3650.pdf. DOI: 10.1007/11681878_14.

[63] F. McSherry and K. Talwar. Mechanism design via differential privacy. Presented at Proceedings of the 48th Annual IEEE Symposium on Foundations of Computer Science. 2007, Available: http://dx.doi.org/10.1109/FOCS.2007.41. DOI: 10.1109/FOCS.2007.41.

[64] F. D. McSherry. Privacy integrated queries: An extensible platform for privacy-preserving data analysis. Presented at Proceedings of the 2009 ACM SIGMOD International Conference on Management of Data. 2009, Available: http://doi.acm.org/10.1145/1559845.1559850. DOI: 10.1145/1559845.1559850.

[65] M. Bezzi. An information theoretic approach for privacy metrics. *Trans.Data Privacy 3(3)*, pp. 199-215. 2010. Available: http://dl.acm.org/citation.cfm?id=2019307.2019309.

[66] P. Mallesh, and A. Roth, "Privacy and Mechanism Design," 2013.

[67] A. Ghosh, T. Roughgarden and M. Sundararajan. Universally utility-maximizing privacy mechanisms. Presented at Proceedings of the Forty-First Annual ACM Symposium on Theory of Computing. 2009, Available: http://doi.acm.org/10.1145/1536414.1536464. DOI: 10.1145/1536414.1536464.

[68] L. Crepin, Y. Demazeau, O. Boissier and F. Jacquenet. Sensitive data transaction in hippocratic multi-agent systems. Engineering Societies in the Agents World IX, A. Artikis, er, G. Picard and L. Vercouter, Eds. 2009, Available: http://dx.doi.org/10.1007/978-3-642-02562-4_5. DOI: 10.1007/978-3-642-02562-4_5.

[69] R. Davis and R. G. Smith. Distributed artificial intelligence. in A. H. Bond and L. Gasser, Eds. 1988, Available: http://dl.acm.org/citation.cfm?id=60204.60230.

[70] P. Lou, Z . Zhou, Y . Chen, J . Fuh, Y. Zhang. Negotiation-Based Task Allocation in an Open Supply Chain Environment. 2006.

[71] R. G. Smith. The Contract Net Protocol: High-Level Communication and Control in a Distributed Problem Solver. Computers, IEEE Transactions *C-29(12),* pp. 1104. 1980. DOI: 10.1109/TC.1980.1675516.

[72] A. Kiayias, B. Yener and M. Yung. Privacy-preserving information markets for computing statistical data. *5628*pp. 32-50. 2009. Available: http://dx.doi.org/10.1007/978-3-642-03549-4_3. DOI: 10.1007/978-3-642-03549-4_3.

[73] J. Gubbi, R. Buyya, S. Marusic and M. Palaniswami. Internet of things (IoT): A vision, architectural elements, and future directions. *Future Gener.Comput.Syst. 29(7),* pp. 1645-1660. 2013. Available: http://dx.doi.org/10.1016/j.future.2013.01.010. DOI: 10.1016/j.future.2013.01.010.

[74] A. J. Jara, P. Lopez, D. Fernandez, J. F. Castillo, M. A. Zamora and A. F. Skarmeta. Mobile digcovery: Discovering and interacting with the world through the internet of things. *Personal Ubiquitous Comput. 18(2),* pp. 323-338. 2014. Available: http://dx.doi.org/10.1007/s00779-013-0648-0. DOI: 10.1007/s00779-013-0648-0.

[75] L. Atzori, A. Iera and G. Morabito. From smart objects to social objects: The next evolutionary step of the internet of things. *IEEE Communications Magazine 52(1),* pp. 97-105. 2014. Available: http://dblp.uni-trier.de/db/journals/cm/cm52.html#AtzoriIM14.

[76] L. Atzori, A. Iera and G. Morabito. The internet of things: A survey. Computer Networks 54(15), pp. 2787. 2010. Available:

http://www.sciencedirect.com/science/article/pii/S1389128610001568. DOI: http://dx.doi.org/10.1016/j.comnet.2010.05.010".

[77] A. Castellani, N. Bui, P. Casari, M. Rossi, Z. Shelby and M. Zorzi. Architecture and protocols for the internet of things: A case study. Presented at Pervasive Computing and Communications Workshops (PERCOM Workshops), 2010 8th IEEE International Conference On. 2010, . DOI: 10.1109/PERCOMW.2010.5470520.

[78] T. Heer, O. Garcia-Morchon, R. Hummen, S. L. Keoh, S. S. Kumar and K. Wehrle. Security challenges in the IP-based internet of things. *Wirel.Pers.Commun. 61(3)*, pp. 527-542. 2011. Available: http://dx.doi.org/10.1007/s11277-011-0385-5. DOI: 10.1007/s11277-011-0385-5.

[79] (March 2012). CIA Chief: We'll Spy on You Through Your Dishwasher. Available: http://www.wired.com/dangerroom/2012/03/petraeus-tv-remote/.

[80] I. Alqassem. Privacy and security requirements framework for the internet of things (IoT). Presented at Companion Proceedings of the 36th International Conference on Software Engineering. 2014, Available: http://doi.acm.org/10.1145/2591062.2591201. DOI: 10.1145/2591062.2591201.

[81] R. H. Weber, Internet of Things – New security and privacy challenges. Computer and Law Security Report. Hong Kong 2010.

[82] (March 19, 2014). Facebook's facial recognition software is now as accurate as the human brain, but what now?. Available: http://www.extremetech.com/extreme/178777-facebooks-facial-recognition-software-is-now-as-accurate-as-the-human-brain-but-what-now.

[83] (March 17, 2014). Facebook Creates Software That Matches Faces Almost as Well as You Do. Available: http://www.technologyreview.com/news/525586/facebook-creates-software-that-matches-faces-almost-as-well-as-you-do/.

[84] Nitu MCSI, MIETE. Configurability in SaaS (software as a service) applications. Proceedings of the 2nd India Software Engineering Conference, February 2009.

[85] M. Armbrust, A. Fox, R. Griffith, A. D. Joseph, R. H. Katz, A. Konwinski, G. Lee, D. A. Patterson, A. Rabkin, I. Stoica and M. Zaharia. Above the clouds: A berkeley view of cloud computing. EECS Department, University of California, Berkeley. 2009Available: http://www.eecs.berkeley.edu/Pubs/TechRpts/2009/EECS-2009-28.html.

[86] K. Busloff, J. Sills, J. Moore. Grid Computing in a Saas Environment. Revionics White Paper, .

[87] M. P. Wellman, W. E. Walsh, P. R. Wurman and J. K. MacKie-Mason. Auction protocols for decentralized scheduling. Games Econ. Behav. pp. 271. 2001. Available: http://www.sciencedirect.com/science/article/pii/S0899825600908224. DOI: http://dx.doi.org/10.1006/game.2000.0822".

[88] T. Vissers, N. Nikiforakis, N. Bielova and W. Joosen. Crying wolf? On the price discrimination of online airline tickets. 2014.

[89] J. Mikians, L. Gyarmati, V. Erramilli and N. Laoutaris. Detecting price and search discrimination on the internet. Presented at Proceedings of the 11th ACM Workshop on Hot Topics in Networks. 2012, Available: http://doi.acm.org/10.1145/2390231.2390245. DOI: 10.1145/2390231.2390245.

[90] A. Hussain, A. Samani and H. Ghennoiwa. A Personal Smart Assistant for Open Environments. International Conference on Artificial Intelligence (ICAI) 2013.

[91] Y. Zhang, H. Ghenniwa and Weiming Shen. Enhancing intelligent user assistance in collaborative design environments. Presented at Computer Supported Cooperative Work in Design, 2005. Proceedings of the Ninth International Conference On. 2005, . DOI: 10.1109/CSCWD.2005.194154.

[92] Y. Zhang, H. Ghenniwa and W. Shen. Agent-based personal assistance in collaborative design environments. pp. 284-293. 2006. Available: http://dx.doi.org/10.1007/11686699_29. DOI: 10.1007/11686699_29.

[93] S. Schiaffino and A. Amandi. Intelligent user profiling. pp. 193-216. 2009. Available: http://dx.doi.org/10.1007/978-3-642-03226-4_11. DOI: 10.1007/978-3-642-03226-4_11.

[94] F. D. McSherry. Privacy integrated queries: An extensible platform for privacy-preserving data analysis. Presented at Proceedings of the 2009 ACM SIGMOD International Conference on Management of Data. 2009, Available: http://doi.acm.org/10.1145/1559845.1559850. DOI: 10.1145/1559845.1559850.

[95] L. Hurwicz and S. Reiter. *Designing Economic Mechanisms* 2006Available: http://books.google.ca/books?id=Mvn8chTLeFwC.

[96] R. K. Dash, N. R. Jennings and D. C. Parkes. Computational-mechanism design: A call to arms. IEEE Intelligent Systems 18(6), pp. 40-47. 2003. Available: http://dx.doi.org/10.1109/MIS.2003.1249168. DOI: 10.1109/MIS.2003.1249168.

Curriculum Vitae

Name: Afshan Samani

Post-secondary Education and Degrees:
Western University
London, Ontario, Canada
2011-2015 PhD.

Related Work Experience
Teaching Assistant / Industrial Research Consultant
Western University
2011-2015

Publications:

Afshan Samani, Hamada H. Ghenniwa and Abdulmotalib Wahaishi (2015), Privacy in Internet of Things: A Model and Protection Framework, ANT 2015.

Abdulmotalib Wahaishi, Afshan Samani and Hamada H. Ghenniwa (2015) SmartHealth and Internet of Things, ICOST 2015.

Afshan Samani, Raafat Aburukba, Adrian Bienkowski and Hamada H.Ghenniwa (2013) Privacy Framework for Open Environments, PASSAT 2013

Afshan Samani and Hamada H. Ghenniwa (2013) Privacy expansion in contract net protocol, Trust in Agent Societies – AAMAS 2013.

Ali. Hussain, Afshan Samani and Hamada H. Ghenniwa (2013), A Personal Smart Assistant for Open Environments ICAI 2013

Afshan Samani, Hamada H. Ghenniwa and Jagath Samarabandu, Risk-based modelling for managing privacy protection, CCECE 2012