Reducing Normative Conflicts in Information Security

Wolter Pieters
Centre for Telematics and Information Technology, University of Twente
P.O. Box 217, 7500 AE Enschede
The Netherlands
w.pieters@utwente.nl

Lizzie Coles-Kemp
Information Security Group, Royal Holloway, University of London
Egham, Surrey TW20 0EX, UK
Lizzie.Coles-Kemp@rhul.ac.uk

ABSTRACT

Security weaknesses often stem from users trying to comply with social expectations rather than following security procedures. Such normative conflicts between security policies and social norms are therefore undesirable from a security perspective. It has been argued that system developers have a "meta-task responsibility", meaning that they have a moral obligation to enable the users of the system they design to cope adequately with their responsibilities. Depending on the situation, this could mean forcing the user to make an "ethical" choice, by "designing out" conflicts. In this paper, we ask the question to what extent it is possible to detect such potential normative conflicts in the design phase of security-sensitive systems, using qualitative research in combination with so-called system models. We then envision how security design might proactively reduce conflict by (a) designing out conflict where possible in the development of policies and systems, and (b) responding to residual and emergent conflict through organisational processes. The approach proposed in this paper is a so-called subcultural approach, where security policies are designed to be culturally sympathetic. Where normative conflicts either cannot be avoided or emerge later, the organisational processes are used to engage with subcultures to encourage communally-mediated control.

Categories and Subject Descriptors

K.6.5 [**Management of Computing and Information Systems**]: Security and Protection

General Terms

Security, Human Factors

Keywords

information security, meta-task responsibility, normative conflicts, policy alignment, security policies, subcultures, system models

NSPW'11, September 12–15, 2011, Marin County, CA, USA.

1. INTRODUCTION

One night, the first author needed to go back to the office. He would need an entrance card to do so. He arrived at the same time as someone else. The first author reasoned as follows. If he opened the door first, the second person might try to tailgate behind him. When the door is open, this becomes very hard to prevent, as he will basically need to block the entrance, which goes against social norms and could lead to him being considered rude and antisocial. Instead, he chose to wait until the second person opened the door. This looked very suspicious to the second person, such that the latter did not want to open the door first either.

In many cases, security rules conflict with cultural practices that are prevalent in the workplace. The philosophy of security policy design is often an adversarial one: one needs to treat the other as a potential enemy, as a means rather than an end, and this goes against our social and ethical predispositions. Social and ethical predispositions are coded into organisational culture, and a family of culture-related drivers also motivate and shape the organisational practices. These drivers include the organisational ideology, beliefs, rituals and myths [29]. Attackers know this, and exploit the resulting weaknesses by means of social engineering. For example, dressing up as Santa Claus seems to work particularly well for accessing restricted areas (so we were told), since people do not regard Santa Claus as a malevolent symbol, are socialised with the ritual of granting access to benign visitors and do not associate Santa Claus with myths of malicious behaviour.

In this paper, we focus on normative conflicts in the security context, i.e., situations in which agents face contradictory expectations on their behaviour, as expressed in organisational or social norms. We evaluate how possible it is to "design out" normative conflicts when designing security policies and systems. We do this from the perspective of security policy alignment (see e.g. [9]). Security policy alignment deals with consistency of security policies, and completeness of the refinement of security policies. Normative conflicts can be seen as a special case of policy inconsistency, namely one where (a) non-security related normative constraints have been explicitly included in the policy model, and (b) conflicts occur in the policies *assigned to humans* (or other normative agents, assuming these exist).

Not all policies that conflict in theory will actually conflict in practice. For example, there may be a policy that forces me to give my manager access to the sales data, and a policy that forbids sharing sales data with family. These would conflict if my manager were my wife. If that is not the

case, there is a conflict in theory, but no conflict in practice. In organisations, there will be many potential conflicts, but not all of them will actually manifest themselves in real-life situations. Models of the organisation (system models) can support such reasoning on actual conflicts, and the interpretation of the system models can be compared across different organisational subcultures.

Rather than preventing attacks directly [33, 24], or discovering problems in real-time [6, 26], our analysis aims at identifying normative policy conflicts in the security policy design phase. Having identified these conflicts, our approach mitigates them by ensuring that security responsibilities are encoded into the design, in terms of (a) the policy itself, (b) security technology, and (c) the organisational processes which implement and maintain the policy. This approach is motivated by the assumption that, even if policy conflicts do not directly cause security threats, they lead to uncertainty with the users, which may in the end weaken their confidence in maintaining security procedures. In this sense, our analysis does not assume a priori that solutions for non-compliance should lie in educating the user; rather, we look for solutions in changing the design of the socio-technical system surrounding the user, which is a composition of technology and the socio-structural system. Whereas technology has implicit (inscribed) constraints on how it can be used, the socio-structural system can be thought of as containing the policies, procedures, organisational processes and controls. Both need to be aligned to allow users to cope adequately with their security responsibilities.

In order for this approach to analysis to succeed, we have to somehow reduce the complexity of a "real life" scenario. Our approach is to take the complex cultural systems within an organisation and reduce the values and rituals to a set of organisational norms, acknowledging that some of the organisational norms will vary from subculture to subculture. It is understood that in reducing cultural systems in this way, some of the complex cultural interaction is abstracted away. As conflicts emerge from interaction between users and the system, and as not all interactions can be predicted, not all conflicts can be prevented. As a result, such an approach will not completely eliminate normative conflict, but will rather reduce it. We therefore aim at providing tools for removing certain conflicts and identifying where others are inevitable. Organisational processes for implementing and maintaining security policies are used to respond to residual and emergent conflict. One can adjust the management processes to compensate (i.e., adapt the awareness training and the focus of the audit to the subcultural requirements), although these are generally weaker in effect than "hard" design measures.

Based on the novel idea of integrating security policies and social norms in a single analysis, our contributions in this paper are the following:

1. An investigation of the fundamental challenges posed by normative conflicts in security (section 2);
2. An extension of reasoning on security policies to include social norms (section 3); and
3. An architecture for a method to reduce normative conflicts in system design (section 4).

Our subcultural approach to security policy design and implementation is based on the following strategy. Firstly, we study in more detail the ideas about embedding responsibility in design, and how these apply to security. Secondly, we identify security policies, ethical norms and personal predispositions, formalise these, and analyse them for potential conflicts. Thirdly, in order to study whether such conflicts can actually occur in the dynamic behaviour of the system, we extend system models with support for both detecting policy conflicts and analysing their contribution to attacks. Finally, we discuss how design changes can support the avoidance of policy conflicts.

2. CHALLENGES OF NORMATIVE CONFLICTS

2.1 Pushing Back Responsibility

If a train driver ignores a red signal and causes an accident, who is responsible? Typically, this would depend on the circumstances (cf. [39]). If the signal was "well-designed", the driver would be held responsible; if the signal was "badly designed", the driver cannot be reasonably assumed to make the right choice, and therefore will not be accountable. The question then becomes what "well-designed" means, and who is responsible for this. Van den Hoven [37] discusses such situations in terms of task responsibility and meta-task responsibility. The train driver then has a task responsibility to avoid accidents, whereas the infrastructure designers have a meta-task responsibility to *enable the driver to take on his responsibility*. In case of confusing signals, for example when it is not clear which track the signal belongs to, the latter is not satisfied.

The example above can be explained in terms of usability: if the user cannot be expected to understand how to work with the system, the meta-task responsibility is not properly fulfilled. Many situations in security have a similar structure, such as when password requirements are too complicated [28]. However, there is an additional problem in security, which is the often required treatment of others as enemies. The particularities of the security domain extend meta-task responsibility from being a sole question of usability to including questions of *reducing policy conflicts for the user*. In particular, these would amount to conflicts between security policies (task responsibilities) and the social and ethical norms that are found within the prevailing organisational subcultures. Strategies for designing out such conflicts from the security design and responding to residual or emergent conflicts is the focus of this paper.

Van den Hoven [38] recently discussed this in the context of preventing moral dilemmas by design. The choice architecture [36] of the system should be such that users are not placed in situations in which it is unclear what they ought to do. Obviously, there is a tradeoff here: some moral dilemmas might be beneficial for the subject, in the sense of training moral sensitivity. We do not follow this ethical discussion here; it suffices to say that in security contexts, one usually does not want the users to make their own moral decisions, but only the *designers* of the policies and systems in place, just as we do not want the train driver to decide that passing a red signal is safe in particular circumstances. The question that we address here, and that enables a substantial practical improvement with respect to the underlying ethical theories, is how normative conflicts can be systematically identified based on the organisational infrastructure,

and how these can be reduced. However, we recognise that design will only reduce the number of dilemmas or the dimensions to existing dilemmas. Therefore, design needs to be understood from the perspective of both developing a policy and the on-going co-creation of the policy as it is deployed and used. It is the position of the authors that moral dilemmas can be partly prevented through the design of security policies, and that the deployment and the management of the policy can be used to reduce the moral dilemmas further.

2.2 Norms and Policies

Norms and policies occur in different forms. In psychology, a distinction is made between injunctive and descriptive social norms [8], where injunctive norms are based on perceptions of approved and disapproved behaviours, whereas descriptive norms are based on perceptions of what others typically *do*. Thus, people may act in certain ways both because they expect their behaviour to be approved, as well as because they perceive this to be common behaviour. In the latter case, there is no expectation of approval. In security, both types of norms may improve or worsen security. For example, if a descriptive norm tells people that their peers provide PIN numbers to machines but not to humans, this provides some resistance against social engineering. Conversely, a descriptive norm stating that others write down their passwords encourages the person concerned to do so as well. For injunctive norms, the expectation that quick delivery of results will be approved, even if security procedures are evaded, can reduce security. This is the common conflict of security versus efficiency norms: people think it is important to get their job done quickly (and they think this will be approved). But changing the perceptions of approval (e.g. by visible disapproval towards others that violate policies) may result in a norm actually enhancing security. Awareness training may play an important role here.

Expectations on what others do influence not only what is considered adequate behaviour for oneself, but also the expected results of choosing a particular behaviour in a situation. Behaviour is not only based on applicable norms, but also on an assessment of the consequences. Thus, in particular when others are involved in achieving the desired outcome, one places trust in these others based on expectations of their behaviour. Trust and norms are intertwined in this way, and the typical patterns of behaviour (and thus descriptive social norms) will depend on how reliable others are in terms of fulfilling expectations. Where security policies are imposed, people may be invited to be less trustful, and this can in turn influence the dynamics of the social norms in place.

Also, norms can be implicit or explicit, where policies are the obvious example of explicitly imposed norms. Analysing normative conflicts requires implicit norms to be made explicit, by qualitative study of the subcultures involved. These subcultures do not only include the users of a system of norms, but also the designers (of the explicit part). Policies imposed in organisations by management are dependent on the cultural norms existing in the management subculture, whereas the effects of the policies manifest themselves in relation to the cultural norms of the users of the systems, who themselves will belong to different subcultures. In practice, the difference between implicit and explicit norms is not clear cut. Norms that are explicit for some may be implicit for others. Again, communication and training can assist in making norms explicit for all, not with the purpose of analysis, but with the purpose of improving security. Communication of this nature often takes place through the individuals who are on the boundaries between one subculture and another, and analysis to identify those individuals can be a valuable activity in a security communication programme. Of course, such explication will only work well if the norms being made explicit are consistent with existing norms of the individuals concerned.

2.3 The Significance of Normative Conflicts in Security Policy Design

The conflict of norms in this way has underpinned a number of significant security breaches in the last five years. In particular, the data loss incident at the UK's HM Revenue and Customs (HMRC) was the subject of much analysis [32]. In this incident, two compact disks (CDs) containing personal details of 25 million people were lost in transit between HM Revenue and Customs (HMRC) and the National Audit Office (NAO). The records on the disks were the details of families claiming child benefit in the UK. Child benefit can be regarded as financial aid for families with children. The data was comprised of address and bank details and was being transferred as part of the external audit of the cases managed under this benefit system. Child benefit is widely used by UK society, and the event caused a public outcry and brought into sharp focus the issues related to how institutions manage citizens' personal data. In response to the data loss, Alistair Darling, the UK's Chancellor of the Exchequer, commissioned a report by UK Chairman at PriceWaterhouseCoopers, Kieran Poynter. The Poynter report specified in detail the events of the incident which culminated in the loss of two CDs. In addition to reporting the events, the report also provided root cause analysis of the issues that led to the incident. In this report, a number of issues relating to policy was presented. These included: the difficulty in interpretation of sufficient authorisation to release the CDs, the inadequate definition of obligations under the policy, poor specification of security controls for data in transit, and the unenforceable nature of the policy regarding the method of CD transport. In this report, two significant issues relating to culture were also identified: below a certain grade, staff would prioritise operational requirements over security requirements, and it had become a routine organisational practice to transfer large amounts of data between HMRC and different external government bodies.

The HMRC example shows that many aspects of compliance with security policies are dependent on cultural interpretation of enforcement and on the interplay between explicit and implicit norms. Various subcultures within HMRC can be identified in the Poynter report. In addition, the Poynter report also shows very clearly the ambiguity that arises between different subcultures and how this ambiguity results in policy misinterpretation. Moreover, the quotations from HMRC staff in the report clearly illustrate that the nature of normative conflict at the organisational level can vary from individual to individual, with each individual having their own interpretation of the security policies. Finally, the HMRC case also indicates that rituals and practices can conflict with the security norms, and tools that help to make these conflicts explicit as they emerge are a necessary addition to security management processes.

2.3.1 Contrasting Cultural Approaches

Traditionally, the implicit stance of information security management approaches is to treat the security policy and the security management organisational processes as structures to be embedded into the organisational culture. Examples of this implicit approach can be found in numerous information security management writings on policy design [4, 27]. This integrative approach is akin to the unitary or strong culture approach described in organisational literature. The philosophy of the strong culture approach is that "effective top managers could build a strongly unified culture by articulating a set of 'corporate' values, perhaps in a vision or mission statement. If those values were reinforced consistently through formal policies, informal norms, stories, rituals, and jargon, in time almost all employees would allegedly share those values." [19, p. 8]. This process of reinforcement is typically part of the process of embedding information security policies within an organisational unit. We see this approach repeatedly in security management strategies, where security culture is integrated with organisational culture through the encoding of security norms within that culture. This integration is driven from "the top" of an organisation (i.e., at senior management level) and is cascaded down through an organisation to organisational unit and individual level.

However, to analyse the HMRC case as resulting from a weak culture where employees had not been appropriately socialised to the "correct" security norms, ignores the fact that the anomalies can be identified as occuring at the points of interaction between subcultural groups within HMRC. Analytically, it is important to consider the presence of subcultures and acknowledge those subcultures when analysing the causes of policy non-compliance. This is because subcultures are not necessarily susceptible to re-shaping and need to be engaged with in a range of ways in order for security policy compliance to be successful. "Generally, the idea of a single organizational level corporate culture, frequently accompanied by the assumption of management being able to shape it, was very popular earlier. (...) Today most scholars emphasize the presence of subcultures in organizations" [3, p. 157]. It follows that if management is unable to shape corporate culture, then the security policy must be subculturally accessible, and cultural change must come about through the interaction between the subcultures. This is a complex set of interactions, not least because different subcultures will have different sets of injunctive and descriptive norms related to security and different security norms will manifest themselves as implicit and explicit.

There are three cultural groupings within HMRC that can be identified in the Poynter report, namely: bureaucratic grouping, which is primarily responsible for the management of the information flow between HMRC and NAO; a production grouping, who processed the information and carried out the transfer tasks; and a professional services grouping, with responsibility for designing the relationship between HMRC and NAO. If the HMRC example is analysed subculturally, a number of key points emerge:

- There were differences between subcultures in terms of prioritising information security policy compliance with operational and budgetary requirements;
- Individuals within the same subculture have different interpretations of accountability for data;
- There are different subcultural approaches to requesting and granting authorisation for data transfer;
- Cultural responses change and are shaped by different episodes in the organisation.

Responding to these subcultural issues through a strong cultural approach is not guaranteed to achieve the desired security compliance. This is because subcultures are naturally emergent and respond to episodes that affect an organisation, and can not be eradicated or predictably changed by education programmes. As Sinclair observed in her study of organisational cultures [34], subcultures exist in organisations not as a creation of cultural management, but emerge through the interaction of individuals. Sinclair presents studies indicating that attempts to implement a "strong culture" are limited in their success, and that a more nuanced interaction with organisational subcultures is necessary in order to design organisational approaches that are successfully embraced by the institution.

The HMRC example also clearly shows that misalignment between policies and the cultures in which they operate can result from not only a mismatch between policies and different subcultural norms, but also from ambiguity that arises between subcultures. This ambiguity is clearly identified in the examples of e-mail exchange between the National Audit Office and HMRC, where the requirements for records transfer are not clearly identifed and are interpreted differently by the different subcultures involved. There is also ambiguity in the presentation and interpretation of the services provided by each subculture. Martin et al. put forward the view that "[w]ithout acknowledging ambiguity, fully, cultural research runs the risk of offering an oversimplified, clearly outlined, cartoonish portrait of a culture that fails to capture the complexity, flux, and contradictions that characterize life in contemporary organizations" [19] (p. 18). The subcultural approach acknowledges the ambiguity by not only considering it in the normative analysis, but also in the design of the security management processes, which include monitoring processes to identify and report the flux and contradictions that can emerge once policy has been designed.

2.3.2 Subcultural vs. Traditional Policy Design

This paper sketches a subcultural approach to designing security policies that seeks to avoid serious compliance failure by identifying where there are conflicts between the security policy and the cultural norms. The response to those normative conflicts either takes place through designing out the conflict, or by adjusting the monitoring processes to identify anomalies and respond where necessary. The subcultural approach contrasts with the typical approach to a serious compliance failure of this nature. The typical approach is to restructure an organisation by aligning it both structurally and culturally in order to support the security policy. This approach has the goal of aligning the organisational culture with the security culture and to adjust the organisational structure by integrating security functions into it. In part, this is what the recommendations in the Poynter report set out to do. The Poynter Report makes 45 recommendations which include the following themes:

- Inclusion of security in the business objectives;
- Introduction of a strengthened organisational hierar-

chy and functions to support information security policy and strategy;

- Integration of the organisational security functions at all levels of the organisation;
- Alignment of day-to-day working life and behaviours with security policies and processes through the process of embedding;
- Stengthening of both the internal and external audit programmes.

In the approach set out in the Poynter report, the security culture and the organisational culture are integrated. The cultures are considered on one organisational level and the security culture is imposed top-down. Also, in this approach non-compliance with policy at departmental level is primarily handled in three ways:

1. Produce detailed procedures at departmental level;
2. Introduce security functions at departmental level;
3. Increase the level of monitoring and compliance evaluation.

The aim of this departmental strategy is to embed security policy at departmental level through guidance and enforcement. The response is based on the assumption that departments require more guidance and a high prioritisation of information security in order to enforce security policy compliance.

In comparison, the subcultural approach would:

- Identify the different subcultures within an organisation;
- Identify where there are different types of normative conflict with the security policy and identify approaches for conflict reduction;
- Specify individual subcultural strategies for implementing security policies;
- Establish subculturally sympathetic methods of promoting information security dialogue between different subcultures.

In the subcultural approach non-compliance with policy at departmental level is handled by identifying the the differences in cultural norms between subgroups, identifying conflicts between cultural norms and the security norms encoded into the policies, and focusing on those individuals who already communicate across subcultures as a conduit for policy and compliance communication. It should be noted that the policy and compliance communication in this model is bidirectional and is used to identify emergent conflicts between policy and culture, identify and respond to ambiguity of interpretation, as well as communicate subcultural security practices.

The subcultural approach recognises that cultural change primarily comes through the dialogue between different subcultures and that different subcultures will engage with security policies in different ways. In this approach too, security policy is implemented bottom-up rather than top-down. The approach is based on the assumption that one of the causes of non-compliance is conflict between the goals of the policy and the organisational culture, and on the view that imposing cultural change top-down has limited effect. The approach is also built on the recognition that there are several cultures within an organisation, and for policy compliance to be successful, policy design needs to recognise these different cultures. The two approaches are not necessarily exclusive of each other, and the subcultural approach could be viewed as a subculturally sympathetic approach to embedding security policies within an organisation.

2.4 Running Examples

In the remainder of the paper, we use the following somewhat simpler examples to illustrate our framework for reducing normative conflicts. The examples consider threats to the organisation as a whole, including the physical and social parts of the system, but the goal is always the security of the information.

Example 1: The master key. A department has an organisational security policy stating that keys shall not be given to others. In case someone needs access to a room that she does not have authorisation for, somebody with authorisation (i.e., with a key) needs to accompany her. In practice, this policy is hard to enforce, because (a) it involves a lot more work for the key owner, conflicting with the individual norm of efficiency, and (b) the policy conflicts with the social norm that colleagues should be trusted. The former is related to the disposition of minimising effort, the latter to the social coherence of the organisation. The secretary therefore often hands over the master key to employees. In the digital world, the key can be thought of as an important password, with slightly different but still quite similar incentives.

Example 2: Tailgating. A department has a policy that only registered employees and visitors can enter the premises. Visitors first need to get a visitor's access card at the reception. However, the entrance doors are on a time delay and stay open for a while after a card is swiped, and they would allow non-registered persons to tailgate behind someone with an access card. The employees have been told not to let this happen, but it occurs regularly anyway. There is a conflict here between the policy, the ethical norm of not treating people as suspicious without a good reason, and the organisational norm of politeness to visitors.

Example 3: Road apple. This example (from [12, 30]) refers to intrapersonal norms rather than social ones, as there are no other human agents involved. The term road apple refers to an apple that is found on a road, tempting the finder to take it. In the IT world, the apple is usually an infected generic dongle, with the logo of the organisation, left by the adversary in a public place in the organisation's premises, such as a canteen. When an employee finds the dongle, she may be tempted to plug the dongle into her laptop [35]. If she does, the dongle will install a rootkit on the hard disk drive without the employee's knowledge.

Many (albeit implicit) examples of reducing normative conflicts by design can already be found in security implementations today. For example, physical means of authentication (such as smartcards) prevent users from trying to help others by providing their username and password over the phone. However, the logic for control selection in examples such as this one is typically risk-based. In the approach

described in this paper, we argue that the identification and analysis of normative conflicts need to be systematically considered as part of any policy evaluation and the traditional risk assessment needs to be contextualised within the relevant normative setting.

2.5 Related Work

The present paper fits into the tradition of both analysing norms and analysing security policies, and, accordingly, the detection of conflicts has been addressed from several angles. In the former area, normative conflicts are seen as (not necessarily security-related) consistency problems in logic or law. In [18], the issue of normative conflicts is discussed from a legal perspective, where normative conflicts may occur between different legal regimes, in particular in international law. In [21] normative conflicts are addressed from the field of deontic logic, i.e., a logic focused on permission and/or obligation. In particular, deontic logics are investigated that tolerate normative conflicts without allowing arbitrary derivations from inconsistencies.

In a security context, [7] and [10] investigate normative conflicts in security policies. The focus is on (1) detecting conflicting permissions and obligations, and (2) what this means for situations where security policies have to be merged.

A third area consists of explanations of how moral reasoning can make people fail to comply with security policies. In this case, the focus is on the effects rather than the origins of conflicts. In this context, [22] focuses on the effects of different types of moral reasoning on compliance.

We find that, although many preventive techniques for information security exist [26, 33, 12], these have so far not been applied to normative conflicts, especially within a setting of organisational norms. Instead, existing research mostly focuses on detecting and resolving conflicts within formally specified security policies. Here, we explicitly focus on the *prevention* of normative conflicts in the design of security architectures and policies, in particular in the context of human behaviour, where both security policies and social or ethical norms may apply. In order to do this, we need to systematically investigate the organisational context in which security policies operate, both in terms of the norms embedded in it, as well as in terms of its role in the actual occurrence of conflicts. The tools developed for preventive analysis of security problems can be adapted for analysing the results of such an investigation, as we will see.

3. REASONING ON SOCIAL NORMS

3.1 Policy Alignment

Security policies may conflict with ethical norms, organisational norms, or personal predispositions. In order to formalise policy conflicts, we need to represent these different types in a single framework. We may even represent malicious/egocentric norms, i.e., norms of users who only aim for their own benefit.

Our perspective in this paper is what is called *policy alignment*. In this perspective, the focus is on whether different policies in an organisation satisfy certain conditions with respect to each other. Here, the concept of policy refers to a rule describing allowed or disallowed sequences of actions. From this perspective, the origin of such rules is not part of the definition of policy. Policies can thus emerge from implicit consensus, from cultural backgrounds, or from explicit agreement on what to enforce. They may even be particular norms held by individuals. Policies may be expressed on a very high level, corresponding to organisational goals, or on a very low level, corresponding to individual actions.

A policy describing allowed action sequences separates the space of action sequences into an "allow" and a "don't care" part (i.e., it is a function from action sequences to {allow, don't care}). The associated requirement is that action sequences in the "allow" part should be made possible. Typically, this refers to action sequences that are essential for the business, and should therefore be allowed. A policy describing disallowed action sequences separates the space of action sequences into a "disallow" and a "don't care" part. The associated requirement is that action sequences in the "disallow" part should be made impossible. Typically, this refers to action sequences that cause security problems, and should therefore be disallowed.[1] In order to make the policies workable, properties can be expressed that cover the notion of policy alignment [9, 11]. A set of policies is *consistent* if there is no action sequence that is covered by both an allow and a disallow policy from the set.

Many real-life constraints of organisations are on a high level. They would, for example, state that certain data should remain within the organisation's boundaries, or that no employee should be able to take away more than a certain amount of money without being detected. In the first example, there are may be many ways in which the data could be taken out, from e-mail to USB sticks. In the second example, the ability to detect depends on the authorisations of employees, but also on having proper monitoring processes in place.

Policies may thus be stated at different levels [11]. At the highest level, there could be policies such as "sales data should not leave the organisation". At lower levels, these may be implemented by physical access control, IT configuration, and clean desk policies. These lower level policies will again separate the space of action sequences in allowed / disallowed and don't care parts. A set of lower level policies can again be tested for consistency. Moreover, we can speak of *completeness* of a set of policies with respect to a set of higher-level policies. A set of policies P_L is *complete* with respect to another set P_H if all action sequences allowed by P_H are allowed by P_L, and all action sequences disallowed by P_H are disallowed by P_L. In other words, the policies in P_L need to be at least as restrictive as the policies in P_H. Thus, the set of action sequences allowed (disallowed) by P_L should be a superset of the set of action sequences allowed (disallowed) by P_H. Completeness of policy refinements with respect to higher-level policies has been dealt with elsewhere [11, 12, 30]. Here, we focus on policy consistency.

Next to organisational policies, cultural norms have their place in organisations. These can be seen as environmental influences with respect to the security system. By explicitly including these norms in the policy model, the consistency analysis can be extended to include such norms. In that case, we could speak of norms representing encouraged and discouraged action sequences. Normative conflicts oc-

[1]It would be possible to add a third type of policy, representing obligatory sequences of actions, but for reasons of simplicity of exposition we do not include those in this first outline of the approach.

	system	environment
high-level	security goals	environmental influences
low-level	business rules	encouraged behaviour

Table 1: High-level and low-level policies

cur when socially encouraged behaviour is disallowed by a security policy. A different type of conflict occurs when socially *discouraged* behaviour is *allowed* by a security policy. The latter type may indicate gaps in organisational policies; we will not discuss this type further here. Therefore, we only focus on *positive* social norms, i.e., norms specifying encouraged behaviour, which may conflict with security policies disallowing this behaviour. In this sense, security policies can be said to be consistent (or complete with respect to higher-level policies) within the context of a certain normative environment, if there are no such conflicts. This is a crucial distinction between the levels within a policy model. Normative analysis needs to recognise the degree to which behaviour can be influenced or controlled at each level.

Norms can be thought of as either duty-based rules, social contracts or rule-utilitarian optimisations, depending on the ethical perspective taken. When seen as rule-utilitarian, meaning an optimisation of rules rather than individual actions for the maximisation of benefit, our view of norms can be interpreted as a rule-utilitarian version of the compliance budget [5], where the budget specifies how much cost people are willing to incur to comply. In this case, rather than speaking about a budget, our perspective focuses on different rules, aimed at optimisation, conflicting in a particular situation. As different optimisation rules conflict, the question of what to do is not immediately decidable. This would then require the agent to re-evaluate the situation in terms of actual benefits rather than simply applying the rules, leading to additional reasoning burden and uncertainty. It is this situation that we wish to prevent to the greatest extent possible.

In Table 1, the relation between high-level and low-level policies, as well as explicit management statements and environmental effects are displayed. Given the technical access topology of an organisation, environmental influences may lead to general encouraged (or discouraged) behaviour. These norms may then conflict with those imposed by information security management. For example, social norms may be in place that lead us to give others what they need to do their job. By contrast, there is an organisational that disallows giving anyone the master key. So how would these two policies interact?

3.2 Identifying Norms

When trying to model interactions between security policies and social norms, these first need to be translated to the same language. In order to formalise organisational and ethical norms, they first need to be identified. Qualitative research methods may be used to elicit organisational and personal norms within an organisational setting. Similarly, security policies may be extracted from documents and/or interviews. All these norms should then be formulated in terms of granting access, whenever this is possible. (This should be possible when the norms are security-relevant.) This is a complex process because norms at the organisational and personal level evolve over time and are influenced by events. This was shown clearly in a study by Pettigrew [29], which used ethnographic research methods to conduct a longitudinal study of organisational culture. This study showed how different events change organisational culture and that, in order to understand present day culture, an understanding of the history of an organisation is required. This observation is echoed in the Poynter report, which points to events that shaped HMRC's cultural expectations about transferring large number of data records to external parties. Therefore, any process to capture norms must be iterative, any analysis must be rooted in the organisational history, and resulting policy design must embrace the fact that organisational and personal norms are not static. Any policy design must thus be part of a wider policy system containing organisational processes that allow for changing policy interpretations.

An iterative process for capturing norms would take such changes into account in each cycle. In summary, the norm identification process underpinning our policy design approach can be described as follows:

1. Obtain formally defined security policies;
2. Use qualitative research methods to discover social and ethical norms at the required level (society, organisation, department, individual);
3. Translate all relevant norms into conditions of granting access: who or what is granting access, to whom or what is access granted, what else needs to be present (credentials).

The second step requires detailed study of the practices at the specified level. Cultural theorist Hofstede [15] identified different measurable dimensions to cultures which enable groupings of different subcultures. Such scales can be adapted and incorporated into technologies, such as monitoring tools, logic formulations or organisational processes. As stated previously, organisation theory typically has an implicit view of organisations as sociocultural systems.

However, in order to conceptualise the interactions between culture and policies at a sufficiently granular level, a more detailed conceptual framework is required. Allaire and Firsirotu [2] proposed a conceptual framework consisting of three interleaving components. Figure 1 illustrates these components. From a security perspective, the cultural system is the values, myths and ideologies that influence perceptions of security and security practice. Whereas, from a security perspective, the sociostructural system is the institutional elements that are used to deploy security: the functions, the policies, the procedures and the processes. The sociocultural system can be enhanced to identify the manner in which the cultural system manifests itself from subculture to subculture.

In today's security management approach, audit focuses on the sociostructural system, i.e., the policies and procedures. In a subcultural approach to policy design, in addition to auditing the policies and procedures, the audit process must be adapted and new audit techniques designed, which are able to identify the cultural system at work in each subculture and monitor for changes to the cultural system. In addition, the audit process must also be able to identify and measure the interaction between the sociostructural system and the cultural system, as it is in the interaction that a normative conflict manifests itself. For example, the cultural myths of a particular subculture may influence a very

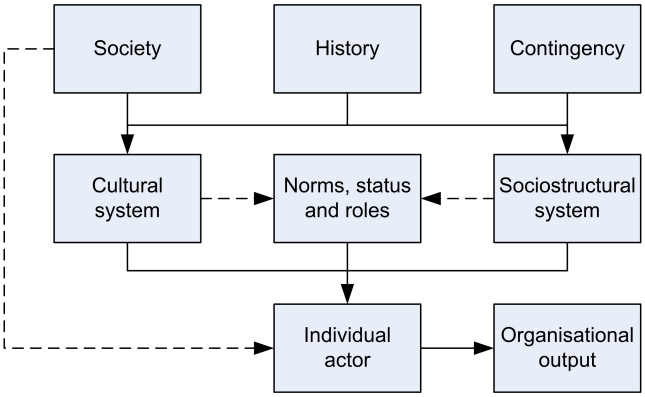

Figure 1: A conceptual framework for organisational culture, simplified from [2, p. 214].

specific interpretation of an aspect of security policy. It may not be until these myths are identified that the cause of particular non-conformances are understood.

The Allaire and Firsirotu framework also identifies the role of the individual and it is important to note that individual actors are also explicitly identified in this conceptual framework.The role of individual interpretations is clearly displayed in the quotations from HMRC staff presented in the Poynter report. Drenth [13] makes the point that an organisation has structures and mechanisms that it uses to standardise its management structures but it also has operational processes which influence how that structure is responded to. In addition, Drenth also makes the point that the type of organisational process deployed is related to the individual attitudes of the workers. The subcultural approach will also require an audit process that can identify differences in implicit and explicit norms as well as separate descriptive from injunctive norms. The audit process can gather this data in a number of ways: from observing security-related behaviour, from analysis of various audit trails as well as through the interviewing of staff.

Thus, studying the organisation from a subcultural perspective will yield a set of norms held by management, subcultures, as well as individuals. From a security perspective, we understand such norms as rules that describe which entities to give access to what, in which situations. This definition excludes norms about actions that do not concern access, such as when to mentally support someone. It does include norms about actions that are not primarily about giving access, but achieve this as a side-effect. For example, when explaining something to someone, norms may regulate which information can and cannot be given over and above what is stipulated in the policy, or they may encourage a more relaxed interpretation of the policy. In the next step, the norms found will be analysed for potential conflicts.

3.3 Discovering Conflicts

Checking the norms for conflicts requires a formalisation step first. To formalise the norms found, we suggest using an organisational ontology, in terms of a type hierarchy of entities. Each norm will then consist of a type of entity granting access, a type of entity receiving access, conditions on the situation, and an allow/deny indication. Inspiration may come from work on formalising ethical norms for artificial agents and robots (see e.g. [23]). Norms may be represented in a formal language, for which XACML [25] might be suitable.

Based on the set of formalised norms, we can now statically check them for potential conflicts, and the corresponding situations (see also [7]). Potential conflicts can occur when for two of the norms in the set:

1. The types of granting entities overlap;
2. The types of receiving entities overlap;
3. There are no contradictory requirements on the situation; and
4. The allow/deny indications of the norms are different.

If for a pair of norms all of the four conditions above apply, there is a potential conflict. The corresponding conflict situation then consists of the intersections of the types (it occurs only where the types overlap), and the union of the situational requirements (the requirements for *both* norms need to be fulfilled).

For example, there may be a policy stating that I should tell my manager about sales data when we are both in the office. There may also be a policy that forbids discussing sales data with anyone in the train. As “my manager” and ”anyone” overlap (and I as a granting entity overlap with myself), clauses 1, 2 and 4 hold. However, “in the office” contradicts with “in the train”, so there is no conflict (unless there is an office in the train, or a train in the office).

For the key example, we can now derive that there could be a possible conflict, between the organisational norm that keys should not be given to others, and the social norm that colleagues should be trusted. The trust norm is first translated to an access norm (give colleagues access to what they need), and then the conflict appears, *when a situation occurs in which the key, a key owner and a colleague are together in the same place.* Note that the conflict only occurs for known colleagues here; other conflicts could extend to the population at large, especially when contact is made by phone or e-mail rather than physical presence (e.g. phishing attacks).

In the road apple example, there is a conflict between organisational policy (don’t plug unknown/unregistered devices into company computers) and individual curiosity (try to view the contents of something using a suitable viewer). As the company computer is also a viewing device for digital contents, whenever an employee, an unregistered data storage device, and a company computer come together, there could be a conflict. Note that the development or selection of a suitable ontology is essential here.

The granularity of the norms is dependent on the level of abstraction we wish to take. Norms may be described on a cultural level (cf. [16]), but it will be more accurate if we describe them on the level of groups or subcultures. We could go as far as to include (assumptions on) individually different norms, based on psychological profiles, but most of the time these will not be necessary for the analysis, and it will be a huge burden in terms of the effort required to gather the data. Thus, there is a tradeoff in terms of the granularity of the models versus the cost of building the model.

When potential conflicts are discovered, these can be fed back into the process of identification of norms. People in

the organisation under study may be confronted with the conflicts, and they may be asked whether or not they recognise the conflict, and how they would deal with it. In this way, the description of the norms could be improved.

4. AN ARCHITECTURE FOR REDUCING CONFLICTS

In the previous section, we have already seen the first steps of a method to reduce normative conflicts in system design, namely (1) identifying norms, and (2) discovering potential conflicts. Such reasoning is only the start of actually reducing the conflicts in design. For this to work, we would also like to be able to identify the situations in which conflicts could actually occur, given the constraints of the organisational infrastructure. Also, we would want to find the attacks that are enabled by such conflicts. For these two purposes we need a dynamic analysis, for which we make use of the paradigm of system models. Finally, we consider how socio-technical designs can be changed to accommodate the results of the analysis.

4.1 Causes and Consequences

System models [12, 30, 33] are means for assessing the vulnerabilities of (socio-technical) systems by finding possible attacks. In the starting configuration, certain connections between entities are assumed, as well as the capabilities of entities. It can then be determined whether evolution of the initial configuration could lead to an undesirable state, i.e., an attack. In this sense, system models allow for automatic generation of attack trees or attack graphs [17, 20], which is already supported by several tools.

Although social engineering may play a role in the attacks found, this has up to now not been expressed in terms of conflicts between different norms. Instead, the behaviour of entities in situation is assumed to be given: they will do everything that is possible. Here, we take one step back, and ask how the behaviour is determined by norms that guide the behaviour. As these norms do not merely state what is possible, but also what should *not* happen, there may be conflicts between them.

We make two adaptations to system models here. Firstly, in addition to identifying attacks, we use system models *to identify normative conflicts* that can actually occur within the evolution of the system. Secondly, we adapt the state transitions in the system model by generating capabilities from norms, rather than assuming static capabilities. This means that the policy conflicts will influence the evolution of the system, and can thereby have a role in the sequences of actions leading attacks (and possibly other policy conflicts).

As a basis, we use the ANKH system model [30], which has the advantage that its concepts and representations are relatively simple (for explanation purposes). When automated analysis is required, a framework with tool support (e.g. [12, 33]) should be chosen. It should not be difficult to translate the system models, or to adapt the automated analysis for the ANKH framework.

In the ANKH framework, socio-technical systems are represented as hypergraphs, where nodes are entities and edges are groups of entities that can interact. The system models are created by gathering information on buildings, IT infrastructure, and access control mechanisms in an organisation. In order to simplify the model, repeated structures

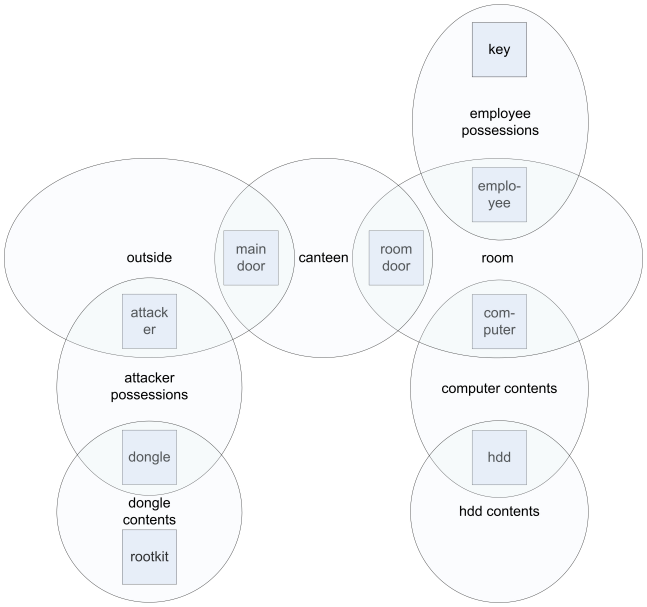

Figure 2: The ANKH model for the initial state of the road apple example (adapted from [30]). "Hdd" stands for hard disk drive, and possessions refer to the items that a person carries with her.

(such as standard rooms) should only be included once in the model. Groups of entities are assigned by determining which entities can directly interact in the starting situation. All entities in a room would be represented by a hyperedge connecting them, as would the entities in the hall. The door, as a member of both groups, controls access in the sense that it can allow an entity to be moved or copied from one to the other. Connecting entities, such as the door in between two rooms, are called *guardians*. Whether a guardian will give access in a particular situation depends on the *capabilities* specified in the model. For example, the door will let someone into the room if she has the right key, if the capabilities state that this is possible. By systematically calculating the possible actions, and relating those to specific goals, the system model can be analysed for potential problems.

The initial state of the system model for the road apple attack is displayed in Figure 2. Capabilities are omitted in the figure. The question we ask the system model is whether the initial state can lead to one of policy conflict, as identified in the previous phase.

Here, we focus on normative conflicts in human entities. The approach may also be applicable to other policy conflicts, for example fire regulations that demand doors to be open in case of fire, versus security policies that require doors to be closed. We will not discuss such conflicts further here.

In order to systematically discover how actual conflict situations can emerge, we supply the situational constraints corresponding to the policy conflict as a goal in system model analysis. The analysis then checks the paths that can lead from an initial configuration to the conflict situation. For example, in the case of the road apple attack, there would be a conflict when an employee, an unregistered data storage device, and a company computer come together. The model can then analyse how such situations can emerge.

In the tailgating example, the feature of the configuration that allows the normative conflict to occur is that, when opening the door, the human herself becomes a second guardian between the open and protected area. The conflict situation thus requires a person in a position where she is able to grant someone else access to the building or room (namely in the doorway holding the door). Whereas access control was meant to be delegated to the door, the human is now also able to grant someone else access. In this guardian position, conflicts between social norms and organisational policies may actually materialise. Obviously, this situation can occur when there is a door that allows a human to act as a guardian between the outside and the inside, i.e., when the door will allow someone *to be both outside and inside*.[2]

In the road apple example, the conflict occurs when the attacker has left the dongle in the canteen, and the employee has picked up the device and moved back to the office (Figure 2). In this situation, both the security policy on connecting unregistered devices as well as the individual curiosity norm apply. Note that one could also consider picking up the dongle as already problematic; one could then analyse a policy conflict there as well.

Based on the existing attack analysis already supported by system models, we can thus analyse system models for the origins of normative conflict situations. To assess the *consequences* of policy conflicts, system models need to be able to reason about the evolution from a policy conflict situation to an attack situation. For this purpose, we need to adapt system models in order to calculate capabilities (i.e., the possible actions) from policy composition, rather than work with pre-defined capabilities.

4.2 Formalising Norm Composition

Characteristic for humans (and potentially other (artificial) agents) is that different (and conflicting) norms may apply in a situation, while still allowing for decisions on actions. The norms somehow need to be composed to derive the resulting capabilities for action in the particular situation. This is not to say that we consider the behaviour to be deterministic, but rather that we need an approximation of likely behaviour to derive the role of normative conflicts in attacks. The question is thus how we can *model* the composition of organisational norms and social norms in human behaviour.

There are several options for deriving the behaviour of an agent in conflict situations. Firstly, one can focus purely on the applicable norms, and enforce a decision in case of conflict by associating the norms with a certain strength. Secondly, one could use a different decision mechanism in case of normative conflicts, e.g. resorting to act utilitarianism (maximising benefits of the action) in case rules do not provide a clear answer. For reasons of manageability of model complexity, we opt for the former solution here.

Assume that in a certain situation n norms apply, $R_1 ... R_n$. Each norm will therefore tell something about whether or not to give entity e access to group G. Assume that the norms will provide guidance in the form of values $v_i \in [0..1]$ associated with norms R_i, where 0 means "don't give access" and 1 means "do give access". Let's assume for now that the norms are deterministic, i.e., $v_i \in \{0, 1\}$. Each norm will in addition have a weight w_i, representing its relative importance from the perspective of the agent.

[2]This is a not-so-trivial aspect to take into account when constructing the system model, but here the focus is on the analysis of which paths can lead to a conflict situation, not on how to develop accurate system models.

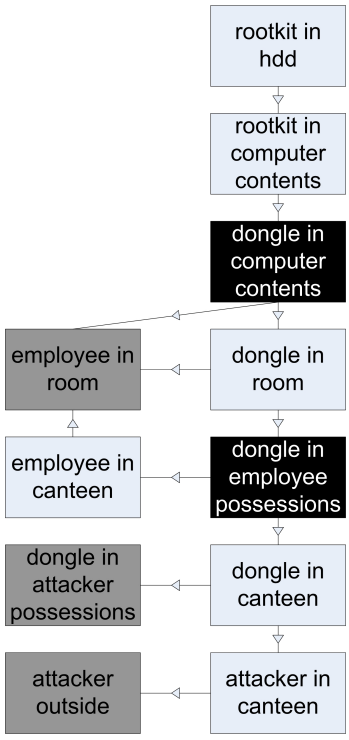

Figure 3: The attack graph constructed for the road apple example (adapted from [30]). The dark grey boxes denote conditions that are already true in the initial state. The black boxes are situations that occur as a result of a policy conflict.

We thus have norms $R_1 ... R_n$ telling the agent to grant entity e access to group G with values $v_1 ... v_n$. The agent will have to decide whether to grant access or not. In this process, some norms may be more important than others. Different strategies for composing such norms can be applied. For example, one may choose the norm with the highest weight and apply it. Alternatively, one could sum the outcomes of the applicable norms and evaluate the total result. In particular, a weighted average could be used to calculate the composed value: $v_c = \frac{\sum_{i=1}^{n} w_i v_i}{n}$.

In the deterministic case, the agent will grant access if $v_c > 0.5$. In the probabilistic case, the values v_i and v_c can be interpreted as likelihoods.

Apart from choosing the most important norm and summing the results, more strategies might be applicable, which could be evaluated in further research. With these capabilities, the system models can derive capabilities from norms, and thereby evolve even in situations of conflicting norms.

We can now follow the evolution of the system model from the conflict situation onwards. Depending on the (likely) capabilities of the agent, we can check which attacks are enabled by the policy conflict. This is done by inputting the conflict situation as the initial state of the model, and violations of high-level policies (e.g. sales data being present outside the organisations) as goal states. By systematically checking the evolution of the system model from the ini-

tial state (see [30], not repeated here), this will yield attack graphs of possible attack routes from the conflict situation to the goal.

Combining the attack graph leading to the conflict situation and the attack graph from the conflict to the goal gives again a standard attack graph. For the road apple example, the complete attack graph, with respect to the goal of the rootkit being in the hard disk drive, is shown in Figure 3 (adapted from [30]). The two conflicts here are in the situation where the employee has to decide whether to pick up the dongle or not, and where the employee has to decide whether to plug it in or not. The conditions making this conflict possible are that there is an item brought in from outside the organisation that can be picked up by the employee. This then leads to a second conflict in the situation where the employee is able to plug the stick into a company computer. Assuming that the individual curiosity norm is stronger than the organisational policy, the possible actions can then lead to an attack. From the point where the stick is put into the computer, an attack is possible by installing the rootkit.

In both situations organisational policies as well as personal preferences may apply, and the resulting behaviour may vary among organisations, groups, and individual employees. If we use the results of the analysis to improve design, models and policies should therefore be improved in an iterative fashion, while the organisation learns about its norms in the process.

4.3 Adapting the Technology

Given that we know which conflicts are possible, how conflict situations can arise, and how conflict situations can lead to attacks, what can we do about it? The traditional solution is awareness training, to make the security policies prevail over any other norms that are in place, but this may not always be effective, especially when social norms are strong. There are, however, other possibilities, in terms of redesigning aspects of the sociocultural system to avoid policy conflicts. Following the steps of our framework, such redesign can be focused on (1) changing norms to eliminate potential conflicts, (2) preventing conflict situations from actually occurring, and (3) preventing conflict situations from leading to attacks.

In the tailgating example, the policy conflict could be avoided by preventing conflict situations. If the doors would only allow one person through at a time, the person would not have to bother thinking about whether or not to close the door in the next person's face. Such doors do actually exist (revolving door or gate that moves only one slot).

In the key example, we could change the capabilities of the key. For example, instead of a regular key, we could use a smartcard that requires biometric identification. In this case, the policy that keys cannot be used by others is inscribed in the key itself. Therefore, the policy can be changed as well: it will no longer be prohibited to give keys away (they will be useless without the owner anyway), or even the applicability of the social norm changes (as the key will now be useless, it doesn't make sense to give it to someone; one would not want to give somebody a useless object).

Designing out normative conflicts by changing the technology may thus take the following forms:

- Preventing the occurrence of *actual* conflicts by not allowing humans to become guardians (tailgating);
- Preventing conflict situations from leading to attacks, by strengthening the credentials, and thereby potentially weakening the norms (master key).

These solutions can be seen as part of a process of security policy refinement [1], where high-level policies are refined into more detailed ones and delegated to policy enforcement points. In this process, the refined policies should be tailored as to minimise the occurrence of normative conflicts, as well as to prevent those from enabling attacks. This can be seen as a process of *alignment* of security policies with organisational norms [11], in which the friction between norms and policies is minimised. In general, security requires minimisation of certain frictions and maximisation of others [14, 31], of which policy alignment is a typical instance.

Where the necessity of avoiding conflicts is highest, technical controls could be used to enforce the elimination of such normative conflicts. However, these approaches are expensive and have implicit usability issues. In other organisational situations it may not be practical or feasible to avoid the conflicts, in which case conflicts might be responded to in the design and implementation of audit and eduction processes, where the objective is to change norms over time, in an attempt to eliminate the potential conflict. The more complex the organisation, the more likely that a blended design approach will take place. Ideally, existing social norms would even be used to improve security, by leveraging them in such a way that they support security policies. This would be a topic for further study.

Thus, based on the suggested analysis of normative conflicts, organisations have different means to respond. The results of such responses can then be checked again in a new iteration of the analysis.

5. CONCLUSIONS AND DISCUSSION

In this paper, we discussed the question how to avoid normative conflicts in information security by design. Normative conflicts are an important source of social engineering attacks, and they may also weaken people's confidence in security procedures, as conflict situations make them unsure what to do. Based on the idea of "pushing back responsibility", these issues should already be addressed in the design phase of security policies and associated artefacts. Our first contribution consists of the analysis of these challenges.

To be able to avoid such conflicts in the design phase, we need a systematic way of testing socio-technical system architectures for such conflicts. Our second contribution is an extension of reasoning on security policies to include social and ethical norms, by providing (a) a method to identify organisational norms, in addition to explicit security policies, and (b) a formalisation of these norms in terms of granting access, and associated consistency tests. As a third contribution, we proposed a framework for analysing and preventing normative conflicts based on existing theories and models of (organisational) security policies. Our framework consists of the above steps (a) and (b) to identify and reason about organisational norms in relation to security policies, as well as (c) an analysis of the causes and consequences of normative conflicts by analysing their role in attacks with system models, and (d) available options to reduce the identified normative conflicts. Step (a), which is

also the most difficult, has been operationalised in terms of a sociocultural study of the organisational context. Phase (b) is based on a formalisation of organisational norms and security policies in terms of access, from which potential conflicts can be identified by statically checking for norms with overlapping applicability but conflicting results. Phase (c) has been operationalised with the use of system models, which can be used to study how conflict situations can occur dynamically, and to which problems they can lead. Our classification of options for preventing normative conflicts in design (d) yields three possibilities, namely (1) changing the norms, (2) preventing conflicts from actually occurring, and (3) preventing conflicts from leading to attacks. The results of such interventions can be fed into a new round of analysis.

To turn this proposal into a full-blown prototype system, implementations of these phases would be needed. This requires further research both on the methodology side and the technical side. In particular, the relation between cultural norms and access norms could be further explored, in terms of a precisely described method for translating one into the other in the context of security. Also, the apparatus for deriving capabilities in system models from norms is an area of future study. It has been suggested that attacks be expressed in terms of difficulty rather than possibility or probability [40], which would enable more fine-grained analysis, but at the same time requires rethinking the composition of norms into capabilities.

The work presented here forms a structured overview of the new area of preventing normative conflicts in security, as well as its challenges, and opens up many possibilities for future study. Our framework allows a focus on policy conflicts next to direct prevention of attack scenarios. This has the additional benefit that people will be placed less often in normative conflict situations, reducing opportunities for social engineering, and potentially increasing people's confidence in what to do. If we really wish to make system security models socio-technical, normative conflicts can therefore not be omitted.

Acknowledgements

The research of the first author is supported by the research program Sentinels (`www.sentinels.nl`). Sentinels is being financed by Technology Foundation STW, the Netherlands Organization for Scientific Research (NWO), and the Dutch Ministry of Economic Affairs. The research of the second author was supported by the Technology Strategy Board; the Engineering and Physical Sciences Research Council and the Economic and Social Research Council [grant number EP/G00255/X].

6. REFERENCES

[1] M. Abrams and D. Bailey. Abstraction and refinement of layered security policy. In M. Abrams, S. Jajodia, and H. Podell, editors, *Information Security: An Integrated Collection of Essays*, pages 126–136. IEEE Computer Society Press, 1995.

[2] Y. Allaire and M. Firsirotu. Theories of organizational culture. *Organization studies*, 5(3):193, 1984.

[3] M. Alvesson. *Understanding Organizational Culture*. Sage, London, 2002.

[4] S. Barman. *Writing Information Security Policies*. New Riders, 2002.

[5] A. Beautement, M. Sasse, and M. Wonham. The compliance budget: managing security behaviour in organisations. In *Proceedings of the 2008 workshop on New security paradigms*, pages 47–58. ACM, 2009.

[6] D. Bolzoni and S. Etalle. Approaches in anomaly-based network intrusion detection systems. In *Intrusion Detection Systems*, volume 38 of *Advances in Information Security*, pages 1–15. Springer, 2008.

[7] L. Cholvy and F. Cuppens. Analyzing consistency of security policies. In *1997 IEEE Symposium on Security and Privacy*. IEEE, 1997.

[8] R. Cialdini. Crafting normative messages to protect the environment. *Current Directions in Psychological Science*, 12:105–109, 2003.

[9] M. Corpuz and P. Barnes. Integrating information security policy management with corporate risk management for strategic alignment. In *Proceedings of the 14th World Multi-Conference on Systemics, Cybernetics and Informatics (WMSCI 2010)*, Orlando, Florida, 29 June - 2 July 2010.

[10] F. Cuppens, L. Cholvy, C. Saurel, and J. Carrere. Merging security policies: analysis of a practical example. In *11th IEEE Computer Security Foundations Workshop*, pages 123–136. IEEE, 1998.

[11] T. Dimkov. *Alignment of organisational security policies: Theory and practice*. PhD thesis, University of Twente, 2012. Forthcoming.

[12] T. Dimkov, W. Pieters, and P. Hartel. Portunes: representing attack scenarios spanning through the physical, digital and social domain. In *Proceedings of the Joint Workshop on Automated Reasoning for Security Protocol Analysis and Issues in the Theory of Security (ARSPA-WITS'10). Revised Selected Papers, Paphos, Cyprus*, volume 6186 of *LNCS*, pages 112–129, Berlin, March 2010. Springer Verlag.

[13] P. Drenth. Culture Consequences in organizations. In P. Drenth, P. Koopman, and B. Wilpert, editors, *Organizational Decision-Making under Different Economic and Political Conditions*, pages 199–206. North-Holland Publishing Co, Amsterdam, 1996.

[14] L. Floridi. The ontological interpretation of informational privacy. *Ethics and Information Technology*, 7:185–200, 2005.

[15] G. Hofstede. Identifying organizational subcultures: an empirical approach. *J. Manage. Stud.*, 35(1):1–12, 1998.

[16] G. Hofstede, C. Jonker, S. Meijer, and T. Verwaart. Modelling trade and trust across cultures. In K. Stølen, W. Winsborough, F. Martinelli, and F. Massacci, editors, *Trust Management: 4th International Conference (iTrust 2006), Proceedings*, volume 3986 of *LNCS*, pages 120–134. Springer, 2006.

[17] D. Kienzle and W. Wulf. A practical approach to security assessment. In *Proceedings of the 1997 workshop on New security paradigms*, pages 5–16, New York, 1998. ACM.

[18] A. Lindroos. Addressing Norm Conflicts in a Fragmented Legal System: The Doctrine of Lex Specialis. *Nordic J. Int'l L.*, 74:27, 2005.

[19] J. Martin, P. Frost, and O. O'Neill. Organizational culture: Beyond struggles for intellectual dominance.

Technical Report 1864, Stanford Graduate School of Business Research Paper Series, 2004.

[20] S. Mauw and M. Oostdijk. Foundations of attack trees. In D. Won and S. Kim, editors, *Proc. 8th Annual International Conference on Information Security and Cryptology, ICISC'05*, volume 3935 of *LNCS*, pages 186–198. Springer, 2006.

[21] C. McGinnis. *Paraconsistency and deontic logic: Formal systems for reasoning with normative conflicts*. PhD thesis, University of Minnesota, 2007.

[22] L. Myyry, M. Siponen, S. Pahnila, T. Vartiainen, and A. Vance. What levels of moral reasoning and values explain adherence to information security rules? an empirical study. *European Journal of Information Systems*, 18:126–139, 2009.

[23] M. Nagenborg, R. Capurro, J. Weber, and C. Pingel. Ethical regulations on robotics in Europe. *AI & Society*, 22(3):349–366, 2008.

[24] V. Nunes Leal Franqueira, R. Lopes, and P. van Eck. Multi-step attack modelling and simulation (MsAMS) framework based on mobile ambients. In *Proceedings of the 24th Annual ACM Symposium on Applied Computing, SAC'2009, Honolulu, Hawaii, USA*, pages 66–73, New York, March 2009. ACM.

[25] O. Open. eXtensible Access Control Markup Language (XACML) version 2.0, 2004.

[26] P. Papadimitriou and H. Garcia-Molina. Data leakage detection. *Knowledge and Data Engineering, IEEE Transactions on*, 23(1):51–63, 2011.

[27] D. Parker. *Fighting Computer Crime*. Wiley, 1998.

[28] S. Parkin, A. van Moorsel, P. Inglesant, and M. Sasse. A stealth approach to usable security: helping IT security managers to identify workable security solutions. In *Proceedings of the 2010 workshop on New security paradigms*, pages 33–50. ACM, 2010.

[29] A. Pettigrew. On studying organizational cultures. *Administrative science quarterly*, 24(4):570–581, 1979.

[30] W. Pieters. Representing humans in system security models: An actor-network approach. *Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications*, 2(1):75–92, 2011.

[31] W. Pieters. The (social) construction of information security. *The Information Society*, 27(5):326–335, September 2011.

[32] K. Poynter. Review of information security at HM Revenue and Customs, June 2008.

[33] C. Probst and R. Hansen. An extensible analysable system model. *Information security technical report*, 13(4):235–246, 2008.

[34] A. Sinclair. Approaches to organisational culture and ethics. *Journal of Business Ethics*, 12(1):63–73, 1993.

[35] S. Stasiukonis. Social engineering the USB way, 2006.

[36] R. Thaler and C. Sunstein. *Nudge: Improving Decisions About Health, Wealth, and Happiness*. Yale University Press, 2008.

[37] M. Van den Hoven. Moral responsibility, public office and information technology. In I. Snellen and W. Van Den Donk, editors, *Public administration in an information age: a handbook*, pages 97–112. IOS, Amsterdam; Washington, DC, 1998.

[38] M. Van den Hoven, G.-J. Lokhorst, and I. Van de Poel. Engineering and the problem of moral overload. *Science and Engineering Ethics*, 2011.

[39] H. Van der Flier and W. Schoonman. Railway signals passed at danger: Situational and personal factors underlying stop signal abuse. *Applied Ergonomics*, 19(2):135–141, 1988.

[40] G. Wyss, J. Clem, J. Darby, K. Dunphy-Guzman, J. Hinton, and K. Mitchiner. Risk-based cost-benefit analysis for security assessment problems. In *Security Technology (ICCST), 2010 IEEE International Carnahan Conference on*, pages 286–295. IEEE, 2010.