Insiders and Insider Threats
An Overview of Definitions and Mitigation Techniques

Jeffrey Hunker
Jeffrey Hunker Associates LLC
hunker@jeffreyhunker.com

Christian W. Probst
Technical University of Denmark
probst@imm.dtu.dk

Abstract

Threats from the inside of an organization's perimeters are a significant problem, since it is difficult to distinguish them from benign activity. In this overview article we discuss defining properties of insiders and insider threats. After presenting definitions of these terms, we go on to discuss a number of approaches from the technological, the sociological, and the socio-technical domain. We draw two main conclusions. Tackling insider threats requires a combination of techniques from the technical, the sociological, and the socio-technical domain, to enable qualified detection of threats, and their mitigation. Another important observation is that the distinction between insiders and outsiders seems to loose significance as IT infrastructure is used in performing insider attacks.

Little real-world data is available about the insider threat [1], yet recognizing when insiders are attempting to do something they should not on a corporate or organizational (computer) system is an important problem in cyber and organizational security in general. This "insider threat" has received considerable attention, and is cited as one of the most serious security problems [2][1]. It is also considered the most difficult problem to deal with because insiders often have information and capabilities not known to external attackers, and as a consequence can cause serious harm. Yet, little real-world data is available about the insider threat.

Especially in the US, there has been substantial research to better understand insider threats and develop more effective approaches. Starting in 1999, RAND conducted a series of workshops to elucidate the necessary research agenda to address this problem [3, 4, 5]. In parallel, the Defense Department produced its own report [6], outlining both a set of policy changes and research directions aimed at addressing the insider threat. Since then, a rich literature studying various aspects of the insider threat problem has emerged.

However, the motivation for work on insider threats appears to differ among countries. Much of the interest in the US seems arguably derives from highly public and damaging national security incidents; Robert Hanssen (arrested in 2001) was an FBI insider who stole and sold secrets to the Russians, and most recently Bradley Manning, a US Army soldier and insider, provided Wiki Leaks with numerous sensitive US government documents. European interest on the other hand appears mostly driven from criminal acts committed by privately employed insiders, as in the $7 billion dollar fraud committed against the French bank Societe Generale by one of its traders, Jerome Kerviel.

Several issues make attacks performed by insiders especially difficult to deal with both from a research and practitioners perspective. There is no uniform or widely accepted definition of either the "insider" or the "insider threat". Indeed, we are forced to conclude that the definition chosen depends on the threat of concern to the specific audience; unfortunately sometimes terminology is used without the precise definition being made clear. Real-world data sets are almost completely missing, a problem shared across cyber security [7], but particularly acute for insider threats. Because by definition the insider is already within at least some element of the organization's security perimeter, security approaches applicable to the "outsider" may not be equally effective for insiders. As a consequence, the insider poses unique security threats arising from his privileged status.

Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications, volume: 2, number: 1, pp. 4-27

[1] The 2008 CSI Computer Crime and Security Survey ranks "insider abuse" second only to viruses in terms of attack types experienced by respondents.

The challenge of satisfactorily resolving these issues distinguishes insider threats from other cyber-security issues. Two distinct schools of possible solution spaces have emerged – *technical* and *sociological/organizational*. The latter examines the psychological aspects and motivation of the insider; that is, why and under what circumstances an insider becomes an insider threat [8], and also the organizational and cultural factors that affect the insider and shape his response to the security environment. Technical approaches use system policy and specifications to prevent, or failing that, identify and minimize the damage done by the threatening insider. Happily, these two different approaches meet in the middle resulting in work on socio-technical approaches.

This separation of technological and sociological concerns is important in security in general, but especially for insider threats, where it allows to separate the technical means for performing and/or mitigating an attack from the sociological means trying to explain the insider's motivation.

This article outlines the principle themes of the insider threat problem and the current thinking of research and practice. We begin in the next section by addressing the foundational challenges of the field, the most fundamental of which is the definition of the insider and the insider threat. We then proceed to outline the solution space (Section 2), considering in turn technical approaches (Section 3), the fused socio-technical approaches (Section 4), and finally the sociological/organizational perspectives (Section 5). We conclude in Section 6 with observations about the many research challenges existing.

1 Foundations

Before presenting the different approaches to dealing with insider threats, we first need to define what we mean with these terms. As we will discuss, both "insider" and "insider threat" are concepts that largely depend on the context in that they are considered, but being able to provide definitions is essential for successful identification and mitigation.

1.1 What is an "Insider"?

There exist many different definitions of the terms "insider" and "insider threat". One common definition is that "an insider is defined as an individual with privileged access to an IT system"[2].

On the surface, this definition seems satisfactory. When machine access was relatively limited and the tasks performed on IT systems were well defined the term "privileged access to an IT system" had a common and well-delineated meaning.

Two developments in recent years have served to confound this picture. One is the now ubiquitous networked computing environment. The other is the increasingly dynamic and porous, if not ill-defined, boundary between the inside of the organization and the outside (consider the range of joint ventures, outsourcing arrangements, consultants and temporary workers in the business world today, for instance).

Today each of the following scenarios could be considered to have "privileged access" and therefore be an insider:

- The recently discharged employee whose system credentials have not yet been revoked (even if organization policy calls for immediate revocation);
- The software developer who designed the organization's systems some time ago and has knowledge of how to access the system;
- A "masquerader" who finds a computer already logged in and uses it without the knowledge of the authenticated user;

[2] This is the definition used by the Department of Homeland Security (US) research project "Human Factors, Awareness, and Insider Threats", 2007-2009.

- The janitor (who has privileged physical access, if not logical access).

As these scenarios' diversity suggests, many different types of insiders exist, and every discussion of the insider seems to start with a different definition – if indeed a definition is provided [1]. Hence clarity and precision in definition is important. We observe that the various definitions of the insider depend on the interpretation and emphasis placed on one or more of the following attributes:

- *Access to the system*, which in turn needs further to distinguish whether or not access is authorized or authenticated(and if authorized, is the authorization legitimate, and legitimate by whom?), and also what form the access takes (*e.g.*, physical versus logical).
- *Ability to represent* the organization to outsiders, as the policies imposed on an actor inside an organization in general are not known to the outside, where the same actor can pretend to be subject to completely different policies. This attribute is often important as a legal basis for defining the insider.
- *Knowledge, e.g.*, the knowledge of the person who originally designed the system, but is not part of the organization or in any way associated with the organization anymore.
- *Trust by the organisation*, empowering an individual. Note that trust has different meanings in security and social sciences, but most definitions of the insider fail to make clear the distinction. In security, trust means dependability and assurance. In the social sciences, trust connotes the 'willingness to be vulnerable based on positive expectations or the actions of others.'

In 2008, a cross-disciplinary workshop on "Countering Insider Threats"' [9] concluded that

> "an insider is a person that has been legitimately empowered with the right to access, represent, or decide about one or more assets of the organization's structure.'

The rationale behind this authority-based definition of the insider is that it removes any specific IT bias from the definition, it focuses on organizational assets rather than a narrow approach based on system credentials, and while people that constitute threats may not be entrusted with access to credentials, they might still have the ability to decide (based on policies) and to represent the organization.

The ability to represent is rather important, as the policies imposed on an actor inside an organization in general are not known to the outside, where the same actor can pretend to be subject to completely different policies – a factor rarely ever being checked.

Specifically, we think that a knowledge-based definition is not sufficiently expressive. Knowledge by an individual (*e.g.*, the knowledge of the person who originally designed the system, but is not part of the organization or in any way associated with the organization anymore) is not a good way of capturing what is an insider, though persons with system or organizational knowledge do represent a threat [10]. However, there are way too many other criteria to consider.

We recommend this definition, having helped to create it, but note that there exist a wide variety of insiders, and their characterization is inherently multidimensional, being defined relative to a particular (computational) framework. In many circumstances it may be best to define what an insider is somewhat loosely, avoiding fine nuances [11].

1.2 What is an "Insider Threat"?

A definition of what an insider threat is obviously depends heavily on the definition of what an insider is. If "an insider is a person that has been legitimately empowered with the right to access, represent, or decide about one or more assets of the organization's structure", what is an insider threat?. One definition is that

> "an insider threat is [posed by] an individual with privileges who misuses them or whose access results in misuse" (alternative definitions can be found, *e.g.*, in [12, 13, 14]).

We would observe, however, that in most circumstances, not every misuse of system privileges by an insider, however defined, is an insider threat of concern, that is, causes real damage.

The term "misuse" raises several problems. "Misuse" implies the presence of rules that specify the conditions of allowable usage; such rules (policies) may in practice be unobserved, non-existent, or poorly written; legal, social and ethical aspects also come into play. It also begs the questions of motivation, consequences, risk to the organization or organizational resources, and the role of the underlying system [5, 15, 16]. Sometimes it can be difficult to separate "acceptable" insider behavior from "unacceptable" behavior. Consider a case where "bad things" happen even though system privileges are not exceeded (*e.g.*, a patient does not receive medical care because a nurse does not/cannot access a medical system) versus where good things happen even though system privileges are exceeded (the nurse uses a password that she is not supposed to know to access the patient's records on the system).

Not surprisingly, several taxonomies for the variety of insider threats have been developed – if one cannot precisely define the problem of question, how can one expect to address it? There exist several examples for taxonomies specifically aimed at insiders [17, 18, 19]. Predd *et al.* [18] define the insider threat relative to characteristics of the individual (motivation, knowledge), the organization (policies, real and official), the system (systems importance in the threat, role of the system in facilitating the threat), and society (ethics and law). This results in a "holistic" or top-down taxonomy. In contrast, Bishop *et al.* [17] define insiders with respect to a resource. Resources are defined as pairs of the resource itself and access privileges, and insiders are defined with respect to these pairs. With this structuring, it is possible to define degrees of insiderness.

However, key factors important as determinants of the insider threat may be difficult to categorize a priori. Knowledge of the insider's intent is desirable but requires all-embracing knowledge. Each determining factor for an insider can be used for defining a taxonomy, for example based on distinctions between:

- *Malicious* and *accidental* threats;
- *Doing something intentionally* (for malice, or good reasons which nonetheless may result in damage) versus events that occur *accidentally*.
- *Obvious* and *stealthy* acts.
- Acts by *masqueraders* (*e.g.*, an individual with a stolen password), *traitors* (malicious legitimate users) and *naive or accidental use* that results in harm.
- *A combination of factors* such as access type; aim or intentionality or reason for misuse; level of technical expertise; and the system consequences of insider threats.

For instance the skill of insiders is a factor in defining the threat posed by malicious insiders, or non-malicious insiders just trying to get their job done. "Motivation" in general is an important question when dealing with insider threats and their consequences. This can cover the whole range from "innocent action", "fun", "technical challenge", "criminal intentions", to "espionage", or a combination of each of these factors. Surprisingly, even though one would expect the contrary, the effect of actions can be equally devastating for each of these motivations. This, of course, makes detecting a threat even more important – but also more complicated.

We would observe that in practice – at least to the extent that we are able to observe real incidents – the problem of real interest is the "real real insider"; an individual deeply embedded in an organization, highly trusted, and in a position to do great damage if so inclined (*e.g.*, a high level executive, or a systems administrator). At the same time it is this kind of insider and the threats he poses that are hardest to deal with [20].

1.3 Issues in Managing the Risk of Insider Threats

A commonly accepted risk management framework and a policy to manage the risk of insider threats do not exist. Risk is defined as (probability of an event) times (consequences of the event). We know little about either.

The impact of insider threats can occur in multiple dimensions: financial loss, disruption to the organization, loss of reputation, and long-term impacts on organizational culture. These impacts can be highly nuanced, and are not well measured or accounted for. For example "bonus round" insider threats (actions taken in anger, revenge, spite regarding bonuses, compensation) can have severe consequences on all levels of an organization. Thus a rather "small" or meaningless motivation can have a rather huge impact. Equally, the impact may not depend on motivation – an innocent act can have as devastating an effect as a maliciously motivated attack. The goal may therefore be to avoid catastrophic consequences regardless of the motivation. These aspects as well as other risk accelerants should be represented in threat models, to acknowledge their importance.

It seems reasonable to assume that the probability of different types of insider threats will vary across organizations and circumstances. Little concrete can be said about this.

Finally, it is unclear how effective various prevention, detection, and response techniques are in reducing the insider threat – and therefore in reducing risk. For instance we lack any clear data to say how effective different security policies are depending on the motivation of the insider. It may be that this does not matter; anecdotally it appears that sometimes it is hard to distinguish the execution and consequences of malicious acts from those due to accidents or naïveté. One the other hand it may matter a great deal. Insiders may legitimately use domains in unexpected ways that might trigger false alarms. Outsiders acting with illegitimately acquired credentials, insiders acting with malicious intent, insiders acting without malicious intent, and accidental behavior are all insider threats, and yet it remains unclear how effective various security policies are against acts stemming from different motives.

1.4 Lack of Data

Little is known about the insider threat [1]. For researchers and practitioners concerned about insider threats perhaps the most fundamental challenge is the lack of real data. Most data about insider threats is anecdotal, and if not anecdotal comes from small, biased data sets. One seminal set of studies conducted by the US Secret Service and Carnegie Mellon analyzed 150 insider crimes incidents with follow up work on 54 cases [21]. The sample is small, focused on only some organization types (those in "critical infrastructures") and most importantly represents those instances where the inside attacker has been caught, prosecuted, and found guilty. Other survey work, notably the Computer Security Institute/FBI annual study [2], are invariably convenience polls (meaning that the respondents provide answers of their own volition) and therefore lack any statistical rigor.

There are good reasons for the lack of data. The absence of good definitions of insider and insider threat confound specifying data requirements. More importantly, most organizations are especially reluctant to share reports of their own insider security problems, for obvious reasons of reputation and possible liability.

Consequently, much of the research on insider threats presupposes a problem (say, masquerader attacks in which an unauthorized user acquires legitimate access rights) and proceeds towards a solution, testing the technique with artificial data sets. One way of creating artificial data is through "capture the flag" exercises, in which study participants are placed in an organizational/computing environment and motivated to act as an insider threat – "capturing the flag" by devising ways of acquiring illicit information [22].

2 Insider Threats Require Multiple Approaches

A study of insider attacks in the banking and finance sector [22] summarized the characteristics of the attacks observed as follows:

- Most incidents required little technical sophistication,
- actions were planned,
- motivation was financial gain,
- acts were committed on the job, and
- incidents were usually detected by non-security personnel and by manual procedures.

Bearing in mind that the sample size was small and biased (perpetrators were caught and punished), this brief list still does suggest both that there is a great deal of room for improvement in countering insider threats, and also that solutions need to engage human systems as well as technical systems.

Because insiders have special knowledge of the organization that outsiders do not have, the interplay between security policies and organizational dynamics is critically important to successfully reducing insider threats. Organizations can and do work around security policies deemed unacceptable, while the real flow of work may be very different than the official representation. Understanding and managing organizational realities is a nuanced and difficult problem in developing effective policies.

This tension between technical approaches and the need to incorporate sociological and organizational insights is a defining aspect of the insider threat field. In the following discussion we acknowledge this tension by examining in turn each of the major solution spaces: technical approaches, approaches that seek to apply socio-technical approaches finally sociological approaches.

In general we consider three major types of insider attacks – misuse of access, bypassing defenses, and access control failure [23]. For each the relative effectiveness of technical versus non-technical approaches varies :

- Misuse of access: the insider has privileges and within those privileges misuses (whatever that means) system resources. This is probably the hardest form of attack to detect or prevent by technical means, since by definition the insider already has legitimate access.
- Bypassing defenses: insiders by definition are already inside the perimeter, and therefore have more opportunity for mischief. Purely technical defenses are insufficient – if they worked, the problem would not exist. Reliance on technical or non-technical detection of anomalous behavior or actual attacks is required.
- Access-control failure: the insider should not have access to specified system resources. This is a technical problem and, while prevention is straightforward, detection of access-control failures is difficult for the same reasons as with access-control misuse [23].

The insider threat is a single name covering a variety of different threats. In practice to date no single approach has proved dominant as a solution. It is important therefore to consider any solution space as likely combining elements of prevention, detection and response against the three canonical attack types, and using both technical and sociological/organizational/psychological components.

3 Technical Approaches

The first area we consider are technical approaches. As mentioned above these include the means for insiders to perform insider threats, as well as means for monitoring and mitigation.

3.1 Policy languages

Policy languages are the tools we can use to express an organization's policies. Ideally there should be a central specification that can be used to generate human-readable descriptions as well as formal representations. The latter ones could be used for analyzing policies for inconsistencies and gaps, the former for documentation and employee education. The biggest concern with policy languages is their granularity and expressiveness. Both need to be chosen at a level that matches the organization and its needs.

An important point to consider with policy languages is their observability and their enforceability. The best policy, especially with respect to insider threats, is one that not only regulates some part of the organisation's workflow, but also describes how to monitor and enforce important aspects.

Formal policy languages offer the benefit of support for identifying contradicting policies. This is especially important when organisations merge and their policies need to be merged as well, or at least need to be adapted.

However, often the best help against an attack is information, not a policy language or a policy – this clearly falls in the socio-technical domain discussed in Section4.

3.2 Access Control

Access control is a way of preventing insider attacks. The ideal access-control policy simultaneously grants the user sufficient privileges to perform necessary tasks, while constraining access according a set of rules. The rules are based on principles of least privilege (the fewer privileges a user is granted the better (in general), escalation (allowing the user to add back certain rights) and separation of duties (essentially splitting actions into separate duties and having multiple persons do each action in order to complete the task) [24].

More formally, access control is the mechanism providing or limiting access to electronic resources based on some set of credentials. This mechanism has two components:

- Authentication, showing who or what you are, *i.e.*, demonstrating possession of certain credentials. Credentialing is often described as presenting some combination of "what you know" (*e.g.*, a password), what you have (*e.g.*, a physical token) and "what you are" (*e.g.*, biometrics). Many also consider a fourth criterion, namely "where you are" (*e.g.*, a certain location).

- Authorization, the system determining if your credentials are sufficient to provide you with a requested type of access [25].

The extent to which explicit, fine-grained access controls can be defined and enforced shapes very directly the type of insider misuse that might occur [26]. In this sense, in a perfect world a perfectly defined and unbreakable access-control system would eliminate insider attacks. In anything less than a perfect world, however, such a system would be unworkable. Resolving this tension between the ideal and the practical defines the loci of research in access control.

In its basic form access control maps users (individual entities or machines or whatever) with access to system resources. Role-Based Access Control (RBAC) is a finer-grained approach. RBAC maps defined "roles" in an organization (rather than users) with access to resources. Users are then assigned roles in order to have access [25]. Temporal RBAC extends this approach by specifying time constraints on when a role can be enabled or disabled [27]. Even more tailored access-control frameworks have been proposed such as a security policy defining subject, object, actions, rights, context and information flow as applicable to the document control domain [27]. Different approaches for implementing access control have also been proposed, an important point being that access control policies also need to be able to specify monitoring and auditing requirements [28].

Access control however has a number of limitations. As noted earlier, even perfect access control will not prevent insider attacks who are only using privileges deemed necessary to get their job done. Studies reveal a disconnect between what "real world" practitioners desire and what the research communities offer [25]. Limiting legitimate access can have a negative effect on the productivity of non-threatening staff.

3.3 Monitoring

Research suggests [29] that harmful insiders operating within their privileges can be identified by observing their patterns of information use, though it is difficult to know what patterns are significant.

In discussions of technical approaches to addressing insider threats, "monitoring" and "detection" are often used interchangeably. Monitoring is a broad term covering various ways of detecting insider attacks or precursors to those attacks, and it has always had a central place in insider threat research.

Monitoring can be conducted in one of three ways: *Misuse detection* identifies defined types of misuse through rule-based detection, in other words matching observed and relevant events against known models of threatening behavior.) This is an extension of intrusion detection systems. Modeling is based on frameworks such as finite state machines, Petri nets, or regular expressions [11]. In theory no false positives should occur but in practice this requires that the modeling framework is sufficiently expressive, and also that the models are specific and up-to-date. Maintaining these models, however, is costly. Also, since only already known attacks can be represented, unknown attacks go undetected. *Anomaly detection* flags significant deviations from expected normal behavior as a proxy for unknown types of anomalous misuse [26]. An anomaly detection system contains three parts:

- A database of information is collected and used to establish behavioral norms. Ideally these norms are adapted as institutional behavior evolves. At design-time the normal behavior is learned from training data, which must not contain any attacks. The machine learning approaches employed to do so are manyfold: statistics, Markov processes, data mining, support vector machines, artificial neural networks, artificial immune systems, etc. [11].
- A monitoring infrastructure is constructed and used to capture events relevant to building dynamic behavioral profiles.
- Actions are triggered when behavior falls outside the expected boundaries, such as signaling a human supervisor to look in on the behavior in question.

Anomaly detection assumes that normal behavior (of a user, a protocol, a process) can be accurately described, that attacks deviate from normal behavior, and that all (or almost all) deviations from normal behavior are attacks [11]. Other sources, however, note that accurately detecting changes in behavior or unusual command sequences is usually not helpful in clarifying the user's intent [22]. The problem of detecting masqueraders (unauthorized users who have illicitly acquired the credentials of an authorized user) is a special variant of anomaly detection and has attracted the attention of some researchers [30], though it is unclear how significant a threat it actually is.

The original work in this technique profiled user command sequences based on statistical features such as the co-occurrence of multiple events [31]. Subsequent work uses a hybrid higher-order Markov chain to identify a signature behavior based on user command sequences [32]. Classification of anomalous events generally uses a naive Bayesian technique. A model of normal user behavior (the self model) is used to generate a probability that a new test block (of behavior) was generated by self. A non-self probability is equivalently generated from a set of training data generated by non-self users. An anomaly score is generated from this model, with an alarm being generated if a threshold value is exceeded [33].

Monitoring can be done at the host level or network layer or both. Host sensors are harder to deploy than network sensors, but some note that "many insider problems do not touch the network level" [22].

Network level monitoring focuses on misuse detection, based on the idea that good evidence of an insider attack exists when an insider accesses information that they do not have a "need to know". ELICIT, one such system, monitors the use of, for example, sensitive search terms, printing to a non-local printer, anomalous browsing activity and retrieving documents outside of the expected social network [29]. Honey-pots and honey tokens are also deployed at the network layer and serve to attract (and trap) ill-intended users with information that the user is not authorized to have or is inappropriate for the user [34].

Finally, there is the issue of how effective monitoring is. Because there is little insider attack data available, it is impossible to tell whether monitoring helps. Monitoring is, reportedly, useful in confirming an already suspected insider attack. There is a controversy as to whether it serves as a deterrent [1]. What evidence there is suggests that monitoring can suffer from both false negatives and false positives. Schonlau *et al.* [35] compared the performance of six masquerade-detection algorithms. All methods had relatively low hit rates and high false alarm rates.

The goal remains to develop dynamic mechanisms for integrating, correlating, and fusing the data sources available on a host in a single anomaly detection system to rapidly detect and identify malicious activities in near real-time, and robust against false positives. Work continues on probabilistic anomaly detection with the goal eventually of modeling user intent.

3.4 Integrated Approaches

Very little research appears to have been done in integrating different approaches and evaluating their effectiveness. Maybury *et al.* [36] developed an integrated detection system using honey-pots, network level sensors, physical security logs, and models of insiders and pre-attack insiders to infer malicious intent. Unfortunately this project was ended before this approach could be evaluated. We await further work in this topic.

3.5 Trusted Systems and other System Hardening

A trusted system ought to be highly resistant to manipulation from within as well as to outside attacks. Early work [37] on this concept included development of the Multics system architecture. The Multics architecture isolated privileged execution domains from less privileged domains and isolated one user from another while still permitting controlled sharing. Sharing was managed via access control lists, access checked dynamic linking, dynamic revocation as well as user independent virtual memory [26]. Work on trusted systems continues. Iyer *et al.* [28] propose an insider threat resistant system using hardware and software based memory randomization with tamper proof key management to secure data from unauthorized intrusions even by the operating system.

3.6 Predictive modeling

Some work has been done to predict threatening insider activity. Altheby proposed a prediction-detection model based on knowledge (that insiders accumulate in their work) and dependencies among different objects and documents pertaining to the organization [38]. Approaches in the same direction aim at modeling systems, and generating attacks from this model of infrastructure, actors, access-control specifications, etc. [39, 40, 41, 42].

4 Socio-Technical Approaches

At the boundary between the just presented technological approaches and sociological approaches there are those techniques that combine both worlds, trying to use finding from both areas to improve detection and mitigation, and to explain success and failures of technological approaches.

4.1 Policies

Policies define the boundaries between permissible and not permissible behavior both on a technical and non-technical level. They not only define proper behavior but implicitly also define the notion of insider. In developing and implementing effective policies for insider threats challenges arise from the inadequacy of the tools available and the complexity of the problems being addressed.

Policy languages

A gap exists between the existing capabilities of policy languages to specify system (and more broadly, organizational) policies, and the needed qualities of policy to adequately prevent insider threats. Policy languages have been developed which are usually quite well suited to support technical issues but are less effective in supporting non-technical aspects of policies.

A major challenge for policy development concerns issues of “context” and “dynamicity”. Stated differently, policies as well as their specification and enforcement are linked tightly to human factors. For example, a given actor might be an insider in one situation, but would be considered an outsider in another situation. Another example are policies that should be obeyed in the general case, but might allow violations in special, emergency cases. In this case it would be the insider’s margin of discretion to decide for or against breaking a policy rule.

However, policy languages lack the capacity to express policies that are “aware” of behavioral aspects and context, thereby being able to handle and regulate abstract events. Domain-specific policy languages really are needed to express contingent policy rules; for example as in real life it would be desirable to develop policy rules under which actions might only be allowed if discretionary circumstances justify their execution. Many systems define the notion of insider relatively to system boundaries.

Policy languages also lack the tools to help users maintain a broader understanding of the systems’ operations and of the subtle effects of different policies. Ideally, policy makers would like some form of capability modeling to understand the impact of a given set of policy changes on the enterprise’s ability to meet its goals [25]; we are far from this capability.

Policy Hierarchy

Once an organization sets a security policy, this might delineate risk-promoting behaviors in the organization and serve as a standard against which to compare behaviors. Policies often are only specified implicitly, and in many organizations we observe that there is a difference between official organizational policy (the de jure policy) and the organizational policy as staff actually implement or understand it (the de facto policy) [18].

More formally, policies themselves are developed based on three sources: 1) legal and regulatory (so-called best practices); 2) business requirements; 3) security requirements. All of these sources can result in implicit or explicit policies, establishing a grey zone where behavior is neither good nor bad.

This potential gap is extended and formalized in the Unifying Policy Hierarchy [43], which established four different levels:

- Oracle Policy. Given perfect knowledge, what would policy be? Deals with inherent vulnerabilities.
- Feasible Policy. With imperfect knowledge, implement oracle as good as possible. Deals with configuration vulnerabilities.
- Configured Policy. What the system implements via configuration. Deals with real time vulnerabilities.
- Real Time Policy. Add in security vulnerabilities.

We observe that both researchers and practitioners often seem to be unaware of the different levels that the same policy may exist in, but instead take for granted that an oracle security policy is provided; given “the” security policy, it is often assumed that this resolves all tensions between organizational culture, work flow, and compliance by (implicitly) enforcing compliance with security practices for the sake of security.

The Policy Hierarchy is useful in highlighting these gaps and conflicts in policies, which appear to be a principle factor in allowing insider threats to occur; in some cases because gaps/conflicts create confusion among insiders in terms of “what is right” or “how do I get my job done”; in other instances because gaps/conflicts create opportunities that malicious insiders can exploit.

For example, consider the consequences of applying security classification to certain resources, an issue especially relevant in military and agency settings. By flagging resources as secret (or in general a high level) to obtain better protection, a conflict is caused in that it becomes immediately evident that the document contains vital information because of implicit information flow based on the classification.

To acknowledge the risk of gaps between policies, we need an analysis of specifications for gaps and conflicts. Reasoning about insider threats it becomes apparent that policies normally do not make explicit who an insider is – an obvious requirement if we want to be able to analyze their hierarchies and fine-tune their impact. If we have policy-language support for specifying the roles of actors, then one may classify certain requests as coming from insiders, or in general build in context-dependent handling of insiders versus outsiders. For example one might want to be able to express that certain requests may only come from insiders, or on an even more context-dependent level, what degree of “insiderness” is required for a certain behavior to be permissive.

Metrics and Measurement

From an organization’s point of view, detection and deterrence must be auditable and traceable to evaluate policies effectiveness. Especially in case of emergency, the only possible action is to audit. Simply measuring the number of incidents and cost of losses is not sufficient to capture the costs and benefits of a particular program of controls and procedures. The concepts of least privilege and separation of duties are excellent guides in discussing insider protection properties that can be measured.

4.2 Monitoring and Profiling

Profiling individual behavior is the basis for anomaly monitoring, and efforts to date have concentrated on relatively straightforward statistical measures [26]. This approach suffers from a number of faults.

For any specific type of insider threat the number of appropriate characteristics that we might want to monitor may be large. As an illustration, in the somewhat similar endeavor of monitoring credit card usage for fraud, up to thirty-five characteristics are monitored to establish legitimate behavior. However, the range of potential misuse by insiders is far greater than the range of credit card misuse, and so effective insider threat modeling may require monitoring a very large number of characteristics.

Statistical modeling also is inadequate for capturing the context or motivation of particular observed actions; multiple motivations may map into a single intent. For instance consider the act (intent) to prop a door open. The motive for this action might be benign (I'm lazy, or am carrying large packages into the room) or malicious (I'm propping the door open to allow unauthorized persons to enter). Observables may capture the intent (opening the door) but not the motivation. This has important implications on the limitations of monitoring, and highlights again the need to establish context for specific actions.

It is unclear how effective monitoring is. Even though monitoring in certain settings has been enhanced significantly, the number of identified incidents has stayed almost constant. At the same time, and even more worrying, cases such as Kerviel and the Liechtenstein case had in common that the attacker intimately knew the monitoring system and knew how to play it.

An important issue for monitoring is the question of how much surveillance is admissible and acceptable in different settings. This sensitivity becomes an increasing concern as monitoring extends beyond technical system observables (*e.g.*, command sequences) towards psychological aspects or non-technical attributes [44].

4.3 Prediction

From the literature about workplace motivations and the psychological profiles of individuals who are or represent potential insider threats, many insider attacks could have been prevented by timely and effective action to address root causes of anger, resentment, or feelings of revenge [45, 46, 47, 48]. Conversely mistaken prediction of potential insider attack could have bad consequences for the individuals under scrutiny as well as ultimately the organization. Many researchers believe in the potential for predictive mechanisms involving psychological indicators to eventually help in avoiding significant asset and information loss serves. Schultz developed a prediction model involving identification of attack-related behaviors and symptoms – indicators that include deliberate markers, meaningful errors, preparatory behaviors, correlated usage patterns, verbal behavior, and personality traits – from which "clues can be pieced together to predict and detect an attack" [49, 50].

As discussed, it is extremely difficult to predict insider attacks [51]. The nature of "trust" and the changeability of human nature (discussed below) contribute to the challenge. Prediction is further complicated by the lack of reliable data. As insider attacks are a sensitive topic, potentially revealing an organization's secrets or damaging its reputation, we have little information about insider attacks, limiting the use of statistical analysis for predicting and recognizing them. Increasing the difficulty of prediction is that even with security mechanisms in place, they will fail every now and then, but it is poorly understood how to detect when they do, and it is not clear either how to best react when they do. In addition, human engagement with technology is emergent and the nature of the emergence is often unpredictable.

Taxonomies[3] of insider threats are the foundation for predictive modeling. Wood [52] was the first to devise a comprehensive taxonomy for describing the insider; the attributes for classification are access,

[3] The following discussion is based on [50].

knowledge, privileges, skills, risks, tactics (attack behaviors), motivation, and process. Each attribute is characterized to simulate malicious insider behavior. Magklaras and Furnell [53] devised a broad taxonomy of misuse that also, unlike Wood, included accidental misuse. Maybury *et al.* [36] noted that insiders left a trail of cyber activity. From the examination of multiple case studies, they devised a taxonomy of cyber events and their associated observables for detecting malicious insider behavior. Cyber-observable data includes network and system activities, information reconnaissance, accesses to assets, manipulation of assets, and data leakage. Other work has applied functional decomposition and graph based modeling to create insider threat models [54, 55].

In system dynamics [8, 56] a method of analyzing the behavior of complex systems over time, simulations are used to model the complexities of the insider threat problem.

The Detection of Threat Behavior project [57] uses a data mining application and a Bayesian network to detect aberrant document access patterns. Aleman-Meza *et al.* [58] examined the use of semantic associations within an ontology for capturing the scope of a cyber investigation and for determining document relevance to an investigation. Magklaras and Furnell [53] have developed an Insider Threat Prediction Model that computes a single-value score from a set of user threat factors such as user sophistication, role, and access to assets. The score is then used to place the user in one of four threat categories: possible internal threat, potential accidental threat, suspicious, or harmless.

Work continues on predicting insider attacks or anticipating potential threats and in methods to use organizational data in addition to cyber data to support the analysis [50].

4.4 Forensics

Forensics are the techniques used during a supposed insider attack or after it, to understand the specifics of what happens or happened. Currently, insider threat forensics appears to be highly undeveloped.

Forensics shares the limitations that monitoring has. Monitoring alone cannot usually discern insider motivation because we do not know or cannot observe the variables we would want. Similarly in forensics the sort of data that we might wish to have had collected might not become apparent until the specifics of the insider incident under review are known – in other words we do not really know in advance what data we might want. While we have decent tools as the result of a large body of work on intrusion detection, it is unclear how these tools help with insider threats.

Current forensics tools often require assumptions such as “only one person had access” or “the owner of the machine is in complete control” . Practical difficulties abound too. Insider attack behavior may be close to the expected behavior. The still often used audit trail can often prove inadequate (redundant, misleading, missing data) – even worse, usually audit trails lack time correlation. Therefore, forensics remains an art, and as an art questions such as what to log or determining the relevance of log data elude clear answers.

4.5 Response to an Insider Attack

From a company point-of-view it turns out to be often preferable to not attempt to stop an ongoing event that is believed to be an insider attack. Consider the impact of false accusations of insider threats on both the individual and the organization. Many suspicious activities which can be observed are correlated with insider threat behavior, but not causally linked. False accusations have multiple deleterious effects: investigative resources are spent, the individuals so accused may quit, seek legal or other recourse (including becoming a real insider threat if they were not before), or be affected psychologically, and the organization’s culture may be affected as well, possibly for extended periods. There is, therefore, a need for decision processes to decide when to intervene and how.

The issue of "response" [4] also applies to cases where predictive modeling or other less formal techniques flag a potential inside attacker. While proactive mitigation – seeking to change behaviors or motivations – may be an advantage to the organization and to the employees who might be helped by timely intervention, it is possible that the practice might be viewed as excessive and invasive by employees. This could exacerbate an already precarious situation and lead to more – or more severe – malicious insider threat events than using less invasive methods alone (which typically do not consider personnel data or at most consider personnel data through traditional management routes). Increased intrusiveness or severity of security measures may contribute to employee job dissatisfaction; management intervention on suspected employee disgruntlement issues may actually increase an employee's frustration level [45]. At the opposite extreme, it is possible that inadequate attention and action can also feed and increase malicious insider activity. One manifestation of this idea has been described as the "trust trap" in which the organization's trust in individuals increases over time, yielding a false sense of security because the trust leads to decreased vigilance toward the threat. This produces fewer discoveries of harmful actions, which in turn increases the organization's level of trust [8, 59]. Here optimistic access control is seen as a viable option, *i.e.*, allowing insider attacks to happen until there is no way back, or even letting them happen unhindered, at the same time ensuring that enough evidence is collected through monitoring. Alternatively it has been suggested that attacks should be allowed to continue, but to "somehow" confine their effects (*e.g.*, by directing the attacks into honey-pots) [26]. It seems often more ruinous to take systems down because of an ongoing attack than to accept the losses and prosecute after the fact.

4.6 More Observations about Policy

Evidence indicates that certain proactive approaches to security management (such as prevention controls, screening, establishing a trust environment) may have benefits over reactive approaches (detect and respond controls, punitive approaches such as sanctions).

In general proactive approaches are needed to complement reactive approaches (that generally help in the short term, but have lower performance in the long term) for effective security in both short and long terms. Theory suggests that attempts to improve the proactive response of an organization fails not because of any inherent deficiency in the techniques themselves, but because of how the introduction of the more proactive techniques interacts with the physical, economic, social and psychological structures in which implementation takes place.

Detection techniques for each of these types of insider threats will vary without any easy generalizations. Technical monitoring, reporting of suspicious behaviors by other staff or outsiders, and detection through non-IT control systems (*e.g.*, financial controls) are the principle detection methods [60]. In large complex systems, risk analysis and focused detection require a significant effort. Not only may monitoring affect trust within an organization, it also is highly nuanced what to look for. For example, an inordinate amount of searching may indicate a masquerade attack (the masquerader is less familiar than the legitimate insider about data structures) – or a forgetful mind.

In summary, fundamentally the attributes key to insider threat identification will be largely context bound. Hence there are always going to be gray areas in how security policies define both insider misuse and proper behavior. Unfortunately while actions are context specific most security polices only inadequately capture the nuances of context.

[4] The following discussion is based on [50].

5 Sociological, Psychological, and Organizational Approaches

Some researchers question the ultimate utility of technological approaches – saying that “the basic element of the insider threat is human: a perpetrator that has abused a position of trust” [1]. Much work coming from the social sciences (sociology, psychology, organizational behavior) has examined the characteristics of insiders and threatening insiders to understand motivations and indicators.

5.1 Insider Threat Motivations

Research [5] characterizing psychological profiles of malicious insiders focuses largely on case studies and interviews of individuals convicted of espionage or sabotage [62, 63, 64]. Band *et al.* [8] and Moore *et al.* [21] summarize findings that reveal behaviors, motivations, and personality disorders associated with insider crimes such as antisocial or narcissistic personality. Anecdotal research is post hoc, mostly derived from interviews with convicted criminals, and speculative in its predictive value. Also, assessing such personality disorders and motivations in an organization is difficult at best, and management or human resources staff may not be able to do so accurately and consistently because a typical organization does not administer psychological or personality inventory tests. Another challenge is that no studies assess and compare the prevalence of these “insider threat” predispositions with occurrence rates in the overall employee population – an important comparison needed to validate the hypothesized relationship.

Most of the work on insider threat motivations has had the goal of developing predictive models that correlate the psychological profiles or behaviors that have been observed in case studies to insider crime – for example, personal predispositions that relate “ to maladaptive reactions to stress, financial and personal needs leading to personal conflicts and rule violations, chronic disgruntlement, strong reactions to organizational sanctions, concealment of rule violations, and a propensity for escalation during work-related conflicts” [8].

Greitzer and Frincke list psychosocial indicators that are considered indications that an individual is a potentially malicious insider [50]. These are based on interviews with human resources managers and others knowledgeable about insider threats. The top five (of twelve) are:

- Disgruntled,
- accepting feedback,
- anger management,
- disengagement,
- disregard for authority, and
- performance issues.

An important assumption is that any such indicator is worthy of consideration only if the employee exhibits extremely serious or grave manifestations of the indicator. In addition, the research conducted to date has found a fair degree of variability in the judgments of human resources experts with regard to the association of these indicators with potential risk of insider attack [61].

[5] This discussion is based on [61]

5.2 Role of Organizational Culture

Most insider threat policies are based on a set of questionable assumptions:

- Once someone is vetted, audit and incident management processes will pick up policy violations.
- Risk assessment processes will pick up changes to individual and group values.
- Training and awareness-building education programs will instill desired security culture.

Security policies to be successful should *support* not *interfere* with organization workflow. In other words, security should support people doing their jobs. Technological security approaches that interfere with work flow may not be accepted and in fact actively subverted by staff (*e.g.,* an iris reader with an "unacceptable" delay before allowing access resulted in staff finding other ways of gaining access).

Compliance with security policies is hard; to make compliance easy for insiders is absolutely necessary for any successful effort to constrain insider threats. Compliance (defined as efforts users will make for purposes they don't understand/agree with) is limited; getting compliance gets more expensive the closer security policies get to the limit of staff tolerance for disruptions to their work flow and social interactions. Fields such as safety and public health have faced similar challenges, and it may be that insider threat policy development should reference other fields for experience in addressing these challenges. Successful security policy needs to demonstrate to insiders the value of security, not just the requirement for security.

Staff security awareness should be considered as sine qua non for a sound insider strategy. Shaw *et al.* [65] describe three levels of user awareness: 1) perception (the user is able to detect threats in the environment), (2) understanding (the user is able to combine information from different sensors, interpret them and use the resulting knowledge to reduce risk in the environment), and (3) prediction (the user is able to predict future attacks and proactively change own behavior to reduce or remove risk).

Key issues in organizational culture: Organizational purpose and management structures affect both security structure and policy. In discussing organizational factors relevant to the insider threat a number of questions must be considered:

- How does trust grow in organizations? In some organizations for example there is lots of trust at the base of the organization but it does not necessarily rise up.
- How can organizations adjust management processes to engender a more positive environment for security? Specifically, how can organizations develop a "reflexive view" that looks at the whole person rather than just as a work resource?
- Whistle blowing: When are organization members comfortable with whistle blowing? Is there a role for technology in extending the whistle blowing capabilities of staff?
- Policy conflict within organizations: It seems reasonable to assume that all organizations have implicit tradeoffs about what is more and less important in their expressions of policy. How can these be made more explicit so that policy and security architectures can more effectively capture these values? Doing so might require a hierarchy of organizational needs like the Maslow hierarchy of individual needs [66].
- Organizational clustering: how much do organizational units cluster in their values? Are there psychological contracts by group clusters within organizations that can be mapped by looking at risk behaviors?

- How can we build robust policy so that when conflicts do arise they can be resolved efficiently in the best interests of the organization?

The role of criminology: The field of criminology can inform insider threat understanding, and within criminology are several theories relevant to insider threats. Earlier theories of deterrence, social bonds, and social learning have been integrated into a theory of planned behavior: for a crime to be committed a person must have both motive and opportunity. As noted, motive matters; for the "cold intellectual attacker" when the possibility of punishment is high and the sanctions are severe potential criminals will be deterred from committing illegal acts, especially when their motives are weak. The goal of "situational" crime prevention is to 1) make the criminal act appear more difficult; 2) make the criminal act more dangerous; 3) reduce the benefit a potential criminal is expecting to recover; and 4) remove the excuses available to the potential malefactor.

5.3 Policies and Human Factors

Policies need to be shaped and evaluated in terms of their human impact. How specific should policies be?

Context of an activity matters a great deal in accurately characterizing abusive insider actions. Context is defined in terms of physical, social, cultural, and temporal dimensions. In adding context into the shaping of security policies, the implication is that there are no standard solutions for workable security – what is usable is what fits. We need security policies appropriate for a specific domain (context) but what happens when insiders use domains in unexpected ways? Security controls must be characterized in the context of the activity, and because there will always be gray areas, those defining security controls must resist the temptation to believe that controls can eradicate all risk. Those defining controls must do this with full participation of management. Those enforcing controls must be willing to accommodate managerial discretion in certain settings.

There is a perception that there are too many policies. The psychological contract with employees generally means that 1) policies need to be made more manageable, and 2) that there is a need to find a way of testing policies to remove redundant policies. The ideal would be a small set of consistent security policies related to behaviors, and fit with business processes, and organizational values and norms.

A critical insight emerging from sociology is the need, if security is to be successful, to link the user community (the insiders in the organization) with the policies being developed and enforced. We perceive that failing to engage staff in security may be the norm but this lack of engagement weakens security.

In other words, security will work best in organizations where people feel that they are part of a larger community. One approach is for organizations to conduct specialized internal exercises [9, 67] with most or all the insiders to identify both the set of useful and acceptable policies, and unique contexts which may result in generalized policies in conflict with organizational needs. Equally it will be key to monitor the implementation of policies "on the ground" by engaging staff and managers on whether policies are appropriate, or interfere with their workflow or culture. Sustained discourse with insiders can help highlight positive examples (of senior executives, for example) and in myth busting; an important goal here is to remove frequently made excuses.

Policies should consider the extent of trust relationships, both formal and tacit, that exist in the organization's work and information flow. Trust is a behavioral expectation, and trust is only necessary in an organization when behaviors cannot be controlled in all dimensions. Trust is also transitive, so one could argue that reducing insider threats would require environments where no one is trusted, or at worst only a few people are trusted; in any event transferred trust relationships should be eliminated.

5.4 Profiling and Monitoring

Observations on “higher level behavior” not observed by systems monitoring may be useful, for example, by using human intelligence to pick up novel attacks and signals that are out of the system. CERT data suggests that in most cases someone else knew about the insider threat actions going on [58], so it is important to find ways to encourage reporting of “suspicious activity” by others. Looking for suspicious (different) behavior patterns by insiders is appealing, but difficult to systematically apply; behavioral patterns include cyber activity, physical movements, physiological signals, and many more. Employment screening data and self/organizational reported data might be useful here, but any screening for behavioral changes is bound to produce false positives from otherwise innocent factors like individual predispositions or lifestyle changes.

While monitoring can help with technical aspects (such as access violations), it does potentially worsen behavioral aspects – in many instances staff resents monitoring. The deciding factor is how much monitoring is acceptable (both ethically and legally), and it is beneficial at all. This question arises at all levels from individual actors, to groups, to companies, to the society as a whole; in many cases there is no consensus as to what is “acceptable”.

5.5 Profiling and Predictive Modeling

Insiders (or people in general) will always act unexpectedly. As one of our colleagues has noted, “there are days when I don’t make sense to myself... let alone to anyone else.” Thus flagging potential insider threats based on departures from “normal” patterns may lack reliability; monitoring for “out of normal” actions may generate too many false positives. There will also always be “gray areas” in drawing the line between insider misuse and proper behavior.

Unpredictability may also be an artefact of creative or dynamic organizations, and as such may be a valuable attribute. Business strategies relying on creativity or dynamism may not work with many forms of recommended security policy – *e.g.*, “the sort of people we hire will not put up with monitoring.”

5.6 Privacy and Legal Considerations

Expectations as to what is “private” or the extent to which users accept restrictions of their use of system resources vary greatly. Issues of security and privacy are founded mostly upon law and ethics. With regard to legal foundations, there is no single source of privacy law in the United States, but rather an incomplete patchwork that includes the U.S. Constitution, the Fourth Amendment to the U.S. Constitution (protecting against unlawful search and seizure), State constitutions, and federal and state statutes [50].

For instance, in most countries significant differences exist in employment conditions – and hence in the capability to monitor or screen staff – between private sector companies and public agencies or the military. Either data protection rights are severely limited when signing a contract with an agency, or there exist concepts such as duty of allegiance, employees are in lifetime service, and they may have knowledge to irredeemable information. Also, some classes of staff – physicians in the US, for instance – anecdotally appear to show marked resistance to monitoring or other restrictions on their use of system resources. For other professionals, like bankers legal requirements to prevent insider trading, for instance, impose on highly compensated professionals significant monitoring and restrictions on system access.

Legal frameworks vary across countries, and in some cases shape the insider threat response even if the laws are not directly concerned with either privacy or security. In the US, for instance, two sector specific laws (HIPPA for health care, and the Financial Services Modernization Act) impose privacy protection requirements on some data, while Sarbanes-Oxley creates an imperative for information security by requiring organizations to certify the accuracy of their financial statements. In practice these laws impose significantly different objectives for an organization’s security policies. Best practices, generally accepted management and organizational guidelines, and standards, both formal and informal, may have regulatory standing or be relevant in legal proceedings.

Very few insider attacks are legally prosecuted. An analysis of US prosecutions revealed slightly more than 100 in the 1995-2008 period. This analysis also showed that US Federal prosecutions focused exclusively on only a small portion of what is believed to be the spectrum of insider attacks [69].

Ethics are important, though exactly how is hard to generalize. For example, if a policy violates an employee's perceived right to privacy, the organization may be at risk of alienating employees or offending clients. The fit between the organization's security policies and "the" ethics of staff and others affected will influence the effectiveness of those policies.

6 Opportunities for Further Research and Development

In every aspect of the insider threat challenge there is a need for much further research and development work to be done [70]:

- Better defining and categorizing the different types of insider threats and attacks is needed as a foundation for collecting and making available to researchers real data on insider attacks. Both needs – definition and categorization, and data collection, are critical;
- Policy languages expressive of conditionality, context, and otherwise equipped to capture the needs of insider threat policy need to be developed;
- There has to be a better and more usable integration of social science and technical perspectives on insider threats in a way that is useful for practitioners;
- There is a need for trustworthy systems that can withstand insider misuse while providing the functionality needed for a wide range of applications, *e.g.*, in business;
- Finer-grained access policies and access controls are needed to help define what constitutes proper usage, thus facilitating the role of insider-misuse detection;
- Ways of evaluating something approaching the "true" effectiveness of different techniques and solutions is needed. Good data is a start; other test and evaluation systems are needed.

7 Conclusion

Work to reduce or "solve" the insider threat takes several forms. Purely technical approaches seek to specify access control, monitor command sequences and other system observables, and harden systems or file structures so as to thwart accidental or malicious activities by insiders. Work in the social sciences uses approaches from psychology, organizational behavior, and sociology to delineate insider threat motivations, attempt to predict insider attacks, and change organizational structures and cultures so as to reduce the motivation while better thwarting insider attacks. Socio-technical approaches combine elements of both perspectives.

This separation of technological and sociological concerns is important in security in general, but especially for insider threats, where it allows to separate the technical means for performing and/or mitigating an attack from the sociological means trying to explain the insider's motivation.

The insider threat is not one problem, but many. The various taxonomies of insider threats developed illustrate the range and subtle ways in which insider attacks and threats vary in important ways between different types. Each organization is different if only because the staff comprising each organization are different.

Modeling human behavior is close to impossible, let alone modeling how it depends on outer and inner factors. A surveillance system is heavily dependent on legal boundaries of what is allowed to be monitored or not, and the amount of data even from legal monitoring can be overwhelming at best. An evaluation system would need to be able to take all the input and models into account, and this is yet another complex task [44].

Fundamentally the problem of insider threats faces two challenges. The insider is already within the system in some way. And the variety of possible bad outcomes is magnified by the interplay between the insider and the organization, not just as a technical but also human system. Consequently no approach has so far offered a satisfactory path towards a solution.

We can omit one closing remark. In times where most attacks in general, and insider attacks in special, are performed using the Internet and organizations' IT infrastructure, the distinction between insiders and outsiders becomes less and less significant. For most organizations it is probably true that the local IT staff has significantly less intimate knowledge of the system than many hackers on the outside do. Even worse, malicious activity from the outside is hard to observe, since contacts exploring the organization's IT landscape can come from many different places at uncorrelated times.

References

[1] C. P. Pfleeger, *Reflections on the Insider Threat*. Springer, 2008, ch. in [71].

[2] R. Richardson, "CSI computer crime & security survey," http://gocsi.com/sites/default/files/uploads/CSIsurvey2008.pdf (last viewed March 2011), 2008.

[3] R. H. Anderson, "Research and development initiatives focused on preventing, detecting, and responding to insider misuse of critical defense information systems," RAND Corporation, Tech. Rep. RAND CF-151-OSD, 1999.

[4] R. H. Anderson, T. Bozek, T. Longstaff, W. Meitzler, M. Skroch, and K. V. Wyk, "Research on mitigating the insider threat to information systems #2," RAND Corporation, Tech. Rep. RAND CF-163-DARPA, 2000.

[5] R. C. Brackney and R. H. Anderson, "Understanding the insider threat," RAND Corporation, Tech. Rep. RAND CF-196-ARDA, 2004.

[6] "DoD Insider Threat Mitigation. Final Report of the Insider Threat Integrated Process Team." US Department of Defense, Office of the Assistant Secretary of Defense (Command, Control, Commuications, and Intelligence). Available from https://acc.dau.mil/CommunityBrowser.aspx?id=37478 (last viewed March 2011), 2000.

[7] C. Gates and C. Taylor, "Challenging the anomaly detection paradigm: a provocative discussion," in *Proc. of New Security Paradigms Workshop 2006 (NSPW'06), Schloss Dagstuhl, Germany*. ACM Press, September 2007, pp. 21–29. [Online]. Available: http://doi.acm.org.globalproxy.cvt.dk/10.1145/1278940.1278945

[8] S. R. Band, D. M. Cappelli, L. F. Fischer, A. P. Moore, E. D. Shaw, and R. F. Trzeciak, "Comparing insider IT sabotage and espionage: A model-based analysis," Carnegie Mellon University, Tech. Rep. CMU/SEI-2006-TR-026, 2006.

[9] C. W. Probst, J. Hunker, M. Bishop, and D. Gollmann, "Countering insider threats," *Dagstuhl Seminar Proceedings*, 2008. [Online]. Available: http://drops.dagstuhl.de/opus/volltexte/2008/1793

[10] C. W. Probst, J. Hunker, D. Gollmann, and M. Bishop, *Aspects of Insider Threats*. Springer, 2010, ch. in [44].

[11] U. Flegel, J. Vayssiere, and G. Bitz, *A State of the Art Survey of Fraud Detection Technology*. Springer, 2010, ch. in [44].

[12] M. Bishop, "The insider problem revisited," in *Proc. of New Security Paradigms Workshop 2005 (NSPW'05), Lake Arrowhead, California, USA*. ACM Press, September 2005, pp. 75–76.

[13] B. Schneier, *Secrets and Lies: Digital Security in a Networked World*. John Wiley & Sons, 2004.

[14] V. Caruso, "Outsourcing information technology and the insider threat," Master's thesis, Air Force Institute of Technology, 2003, available from http://handle.dtic.mil/100.2/ADA415113 (last viewed March 2011).

[15] N. Einwechter, "Preventing and detecting insider attacks using ids," Available from http://www.symantec.com/connect/articles/preventing-and-detecting-insider-attacks-using-ids (last viewed March 2011), 2002.

[16] E. E. Schultz and R. Shumway, *Incident Response: A Strategic Guide to Handling System and Network Security Breaches*. New Riders, 2001.

[17] M. Bishop, S. Engle, S. Peisert, S. Whalen, and C. Gates, "Case studies of an insider framework," in *Proc. of the 42nd Hawaii International Conference on System Sciences (HICSS'09), Hawaii, USA*. IEEE, January 2009, pp. 1–10.

[18] J. Predd, S. L. Pfleeger, J. Hunker, and C. Bulford, "Insiders behaving badly," *IEEE Security and Privacy*, vol. 6, pp. 66–70, July 2008. [Online]. Available: http://portal.acm.org/citation.cfm?id=1441365.1441416

[19] G. B. Magklaras and S. M. Furnell, "Insider threat prediction tool: Evaluating the probability of it misuse," *Computers & Security*, vol. 21, no. 1, pp. 62 – 73, 2001. [Online]. Available: http://www.sciencedirect.com/science/article/B6V8G-452D9TY-C/2/d3ce0be409d1fbb34d981e7f0cfecc13

[20] C. W. Probst and J. Hunker, "The risk of risk analysis and its relation to the economics of insider threats," in *Economics of Information Security and Privacy*, T. Moore, D. Pym, and C. Ioannidis, Eds. Springer, 2010, pp. 279–299.

[21] A. P. Moore, D. M. Cappelli, and R. F. Trzeciak, *The 'Big' Picture' of IT Sabotage Across U.S. Critical Infrastructures*. Springer, 2008, ch. in [71].

[22] M. B. Salem, S. Hershkop, and S. J. Stolfo, *A Survey of Insider Attack Detection Research*. Springer, 2008, ch. in [71].

[23] S. M. Bellovin, *The Insider Attack Problem: Nature and Scope*. Springer, 2008, ch. in [71].

[24] S. Smith and J. Marchesini, *The Craft of System Security*. Addison-Wesley Professional, 2007.

[25] S. Sinclair and S. W. Smith, *Preventative Directions for Insider Threat Mitigation via Access Control*. Springer, 2008, ch. in [71].

[26] P. G. Neumann, *Combating Insider Threats*. Springer, 2010, ch. in [44].

[27] S. Pramanik, V. Sankaranarayanan, and S. Upadhyaya, "Security policies to mitigate insider threat in the document control domain," in *Proc of the 20th Annual Computer Security Applications Conference (ACSAC'04), Tucson, Arizona, USA*. IEEE, December 2004, pp. 304–313.

[28] R. K. Iyer, P. Dabrowski, N. Nakka, and Z. Kalbarczyck, *Preconfiguarable Tamper-resistant Hardware Support Against Insider Threats: The Tested ILLIAC Approach*. Springer, 2008, ch. in [71].

[29] M. A. Maloof and G. D. Stephens, "Elicit: a system for detecting insiders who violate need-to-know," in *Proc. of the 10th International Conference on Recent Advances in Intrusion Detection (RAID'07), Gold Goast, Australia, LNCS*, vol. 4637. Springer-Verlag, August 2007, pp. 146–166.

[30] B. M. Bowen, M. B. Salem, A. D. Keromytis, and S. J. Stolfo, *Monitoring Technologies for Mitigating Insider Threats*. Springer, 2010, ch. in [44].

[31] B. D. Davison and H. Hirsh, "Predicting sequences of user actions," in *Predicting the Future: AI Approaches to Time Series Problems*. AAAI Press, 1998, pp. 5–12.

[32] W. Ju and Y. Vardun, "A hybrid high-order markov chain model for computer intrusion detection," National Institute of Statistical Sciences, Tech. Rep. 92, 1999.

[33] K. S. Killourhy and R. A. Maxion, *Naive Bayes as a Masquerade Detector: Addressing a Chronic Failure*. Springer, 2008, ch. in [71].

[34] L. Spitzner, "Honeypots: Catching the insider threat," in *Proc. of the 19th Annual Computer Security Applications Conference (CSAC'03), Las Vegas, NV, USA*. IEEE, December 2003, pp. 170–179.

[35] M. Schonlau, W. DuMouchel, W. H. Ju, A. F. Karr, M. Theus, and Y. Vardi, "Computer Intrusion: Detecting Masquerades," *Statistical Science*, vol. 16, no. 1, pp. 58–74, 2001. [Online]. Available: http://dx.doi.org/10.2307/2676780

[36] M. Maybury, P. Chase, B. Cheikes, D. Brackney, S. Matzner, T. Hetherington, B. Wood, C. Sibley, J. Marin, T. Longstaff, L. Spitzner, J. Haile, J. Copeland, and S. Lewandowski, "Analysis and detection of malicious

insiders," The MITRE Corporation, Tech. Rep., 2005.

[37] J. H. Saltzer, "Protection and the control of information sharing in multics," *Commun. ACM*, vol. 17, pp. 388–402, July 1974. [Online]. Available: http://doi.acm.org/10.1145/361011.361067

[38] Q. Althebyan and B. Panda, "A knowledge-base model for insider threat prediction," in *Proc. of the 2007 IEEE SMC Information Assurance and Security Workshop (IAW '07), West Point, New York, USA*. IEEE, June 2007, pp. 239 –246.

[39] T. Dimkov, W. Pieters, and P. Hartel, "Portunes: generating attack scenarios by finding inconsistencies between security policies in the physical, digital and social domain," University of Twente, Tech. Rep., 2009.

[40] T. Dimkov, W. Pieters and P. Hartel, "Portunes: representing attack scenarios spanning through the physical, digital and social domain," in *Proc. of the Joint Workshop on Automated Reasoning for Security Protocol Analysis and Issues in the Theory of Security (ARSPA-WITS'10), Paphos, Cyprus, LNCS*. Springer Verlag, March 2010.

[41] C. W. Probst and R. R. Hansen, "Analysing access control specifications," *2009 Fourth IEEE International Workshop on Systematic Approaches to Digital Forensic Engineering*, pp. 5,6, 2009.

[42] ——, "An extensible analysable system model," *Information Security Technical Report*, vol. 13, pp. 235–246, 2008.

[43] M. Bishop, S. Engle, S. Peisert, S. Whalen, and C. Gates, "We have met the enemy and he is us," in *Proc. of the 2008 Workshop on New Security Paradigms (NSPW'08), Lake Tahoe, California, USA*. ACM Press, September 2008, pp. 1–12.

[44] C. W. Probst, J. Hunker, D. Gollmann, and M. Bishop, Eds., *Insider Threats in Cybersecurity*. Springer, 2010.

[45] E. D. Shaw and L. F. Fischer, "Ten tales of betrayal: The threat to corporate infrastructures by information technology insiders analysis and observations," Defense Personnel Security Research Center (PERSEREC), Monterey, CA, Tech. Rep. 05-13, 2005.

[46] K. Aquino, T. M. Tripp, and R. J. Bies, "Getting even or moving on? power, procedural justice and types of offense as predictors of revenge, forgiveness, reconciliation and avoidance in organizations," *Journal of Applied Psychology*, vol. 91, no. 3, pp. 653–668, 2006.

[47] K. Aquino, T. M. Tripp, and R. J. Bies, "How employees respond to personal offense: The effects of blame attribution, victim status, and offender status on revenge and reconciliation in the workplace," *Journal of Applied Psychology*, vol. 86, no. 1, pp. 52–59, 2001.

[48] D. De Cremer, "Unfair treatment and revenge taking: The roles of collective identification and feelings of disappointment," *Group Dynamics: Theory, Research and Practice*, vol. 10, no. 3, pp. 220–232, 2006.

[49] E. E. Schultz, "A framework for understanding and predicting insider attacks," *Computers & Security*, vol. 21, no. 6, pp. 526 – 531, 2002. [Online]. Available: http://www.sciencedirect.com/science/article/B6V8G-46XGM6D-9/2/69a602ad8a9c42af84570b33777c4c0c

[50] F. L. Greitzer and D. A. Frinke, *Toward Predictive Modeling for Insider Threat Mitigation*. Springer, 2010, ch. in [44].

[51] L. A. Kramer, J. Richards J. Heuer, and K. S. Crawford, "Technological, social and economic trends that are increasing u.s. vulnerability to insider espionage," Defense Personnel Security Research Center (PERSEREC), Monterey, CA, Tech. Rep. 05-10, 2005.

[52] B. J. Wood, "An insider threat model for adversary simulation," in *[4]*, 2000.

[53] G. B. Magklaras and S. M. Furnell, "Insider threat prediction tool: Evaluating the probability of it misuse," *Computers & Security*, vol. 21, no. 1, pp. 62 – 73, 2001. [Online]. Available: http://www.sciencedirect.com/science/article/B6V8G-452D9TY-C/2/d3ce0be409d1fbb34d981e7f0cfecc13

[54] J. W. Butts, R. F. Mills, and R. O. Baldwin, "Developing an insider threat model using functional decomposition," in *Computer Network Security*, ser. Lecture Notes in Computer Science, V. Gorodetsky, I. Kotenko, and V. Skormin, Eds. Springer Berlin / Heidelberg, 2005, vol. 3685, pp. 412–417. [Online]. Available: http://dx.doi.org/10.1007/11560326 32

[55] R. Chinchani, A. Iyer, H. Q. Ngo, and S. Upadhyaya, "Towards a theory of insider threat assessment," in *Proc. of the 2005 International Conference on Dependable Systems and Networks (DSN'05), Yokohama, Japan*. IEEE, June-July 2005, pp. 108–117.

[56] D. M. Cappelli, A. G. Desai, A. P. Moore, T. J. Shimeall, E. A. Weaver, and B. J. Willke, "Management and education of the risk of insider threat (merit): System dynamics modeling of computer system sabotage," in *Proc. of the 24th International Conference of the System Dynamics Society, Nijmegen, The Netherlands*. System Dynamics Society, July 2006.

[57] P. C. Costa, D. Barbará, K. B. Laskey, E. J. Wright, G. Alghamdi, S. Mirza, M. Revankar, and T. Shackelford, "Dtb project: A behavioral model for detecting insider threats," in *Proc. of the 2005 International Conference on Intelligence Analysis, McLean, VA, USA*. The MITRE Corporation, 2005.

[58] B. Aleman-Meza, P. Burns, M. Eavenson, D. Palaniswami, and A. Sheth, "An ontological approach to the document access problem of insider threat," in *Intelligence and Security Informatics*, ser. Lecture Notes in Computer Science, P. Kantor, G. Muresan, F. Roberts, D. D. Zeng, F.-Y. Wang, H. Chen, and R. C. Merkle, Eds. Springer Berlin / Heidelberg, 2005, vol. 3495, pp. 45–47. [Online]. Available: http://dx.doi.org/10.1007/11427995_47

[59] M. Keeney, E. Kowalski, D. Cappelli, A. Moore, T. Shimeall, and S. Rogers, "Insider threat study: Computer system sabotage in critical infrastructure sectors," U.S. Secret Service and CERT Coordination Center, Tech. Rep., 2005, available from http://www.cert.org/insider_threat/insidercross.html (last viewed March 2011).

[60] C. Hlavica, U. Klapproth, and F. M. Hülsberg, Eds., *Tax Fraud & Forensic Accounting*. Gabler, 2011.

[61] M. Bishop, S. Engle, D. A. Frincke, C. Gates, F. L. Greitzer, S. Peisert, and S. Whalen, *A Risk Management Approach to the "Insider Threat"*. Springer, 2010, ch. in [44].

[62] Director of Central Intelligence/Intelligence, "Community staff memorandum ics 0858-90: Project slammer interim report (u)," Central Intelligence Agency, 1990, project Slammer is a CIA-sponsored study of Americans convicted of espionage against the United States. A declassified interim report is available at: http://antipolygraph.org/documents/slammer-12-04-1990.shtml and http://antipolygraph.org/documents/slammer-12-04-1990.pdf.

[63] M. G. Gelles, "Exploring the mind of the spy," In: *Employees' guide to security responsibilities: Treason 101*, Texas A&M University Research Foundation, 2005.

[64] J. Krofcheck and M. G. Gelles, *Behavioral consultation in personnel security: Training and reference manual for personnel security professionals*, Yarrow Associates, 2006.

[65] E. D. Shaw, J. M. Post, and K. G. Ruby, "Inside the mind of the insider," *Security Management*, vol. 43, no. 12, p. 34, 1999.

[66] A. H. Maslow, "A theory of human motivation," *Psychological Review*, vol. 50, no. 4, pp. 370–396, 1943.

[67] C. W. Probst, J. Hunker, M. Bishop, L. Coles-Kemp, and D. Gollmann, "Insider threats: Strategies for prevention, mitigation, and response," *Dagstuhl Seminar Proceedings*, 2010. [Online]. Available: http://drops.dagstuhl.de/opus/volltexte/2010/2903

[68] M. R. Randazzo, M. M. Keeney, E. Kowalski, D. Cappelli, and A. Moore, "Insider threat study: Illicit cyber activity in the banking and finance sector," U.S. Secret Service and CERT Coordination Center, Tech. Rep., 2004, available from http://www.cert.org/archive/pdf/bankfin040820.pdf (last viewed March 2011).

[69] J. Hunker and C. Bulford, "Federal prosecution of insider threats demonstrates need for reform: Analysis based on data base of federal prosecutions since 1995," 2009.

[70] D. Maughan et al., *A roadmap for cybersecurity research*, Department of Homeland Security, 2009.

[71] S. Stolfo, S. Bellovin, S. Hershkop, A. Keromytis, S. Sinclair, and S. W. Smith, Eds., *Insider Attack and Cyber Security: Beyond the Hacker*. Springer, 2008.

Christian W Probst is an Associate Professor in the department for Informatics and Mathematical Modelling at the Technical University of Denmark, where he works in the section for Language-Based Technologies. The motivation behind Christian's research is to realize systems with guaranteed properties. An important aspect of his work are questions related to safety and security properties, most notably insider threats. He is the creator of ExASyM, the extendable, analysable system model, which supports the identification of insider threats in organisations.

Jeffrey Hunker is a Principal of Jeffrey Hunker Associates LLC, a research and advisory firm specializing in information security and national security issues. Jeffrey's research and advisory work focuses on the development and evaluation of public policy, national security policy, and corporate strategies to problems of information assurance and cyber security. Of particular focus are considerations of risk management and the economics of cyber security, including development of insurance and risk shifting mechanisms, and the refinement of evaluative metrics for public policies and corporate strategies.

Together, Jeffrey Hunker and Christian W Probst have co-organized cross-disciplinary workshops on insider threats, and have co-edited a book on the topic.