Countering Insider Threats
Dagstuhl Seminar 08302
July 20st to July 25th 2008

Matt Bishop[1], Dieter Gollmann[2], Jeffrey Hunker[3], Christian W. Probst[4]

[1] University of California, Davis
bishop@cs.ucdavis.edu
[2] Technical University Hamburg-Harburg
diego@tu-harburg.de
[3] Carnegie Mellon University
jhunker@andrew.cmu.edu
[4] Technical University of Denmark, Informatics and Mathematical Modelling
probst@imm.dtu.dk

Abstract. This article summarizes the objectives and structure of a seminar with the same title, held from July 20th to July 25th, 2008, at Schloss Dagstuhl, Germany. The seminar brought together researchers and policy-makers from all involved communities, to clarify what it is that identifies an insider threat, and to develop a common vision of how an insider can be categorized as well as an integrated approach that allows a qualitative reasoning about the threat and the possibilities of attacks. This report gives an overview of the discussions and presentations during the week, as well as the outcome of these discussions.

1 Introduction

The "insider threat" or "insider problem" has received considerable attention, and is cited as the most serious security problem in many studies. It is also considered the most difficult problem to deal with, because an "insider" has information and capabilities not known to other, external attackers. However, the term "insider threat" is usually either not defined at all, or defined nebulously.

The difficulty in handling the insider threat is reasonable under those circumstances; if one cannot define a problem precisely, how can one approach a solution, let alone know when the problem is solved? It is noteworthy that, despite this imponderability, definitions of the insider threat still have some common elements. For example, a workshop report defined the problem as malevolent (or possibly inadvertent) actions by an already trusted person with access to sensitive information and information systems. Elsewhere, that same report defined an insider as someone with access, privilege, or knowledge of information systems and services. Another report implicitly defined an insider as anyone operating inside the security perimeter—while already the assumption of only having a single security perimeter may be optimistic.

The goal of this Dagstuhl seminar was to bring together researchers and practitioners from different communities to discuss in a multi-national setting what

Dagstuhl Seminar Proceedings 08302
Countering Insider Threats
http://drops.dagstuhl.de/opus/volltexte/2008/1793

the problems are we care about, what our response is, which factors influence the cost of dealing with insider threats and attacks, and so on. In a time where we barely understand which factors cause insider threats, and our solutions are scattered all over communities, areas, and instruments, this coordinated action between the involved communities seems to be needed more than ever.

This Dagstuhl seminar was, to our knowledge, the first European seminar focusing on insider threats bringing together US and European researchers and practitioners. The five days of the seminar allowed not only for a rich assortment of presentations, but even more importantly for extended discussions, both formal and informal, among the participants. We even had the opportunity for a structured exercise that challenged participants to define specific insider threats, develop the appropriate responses, and critique each others problem-solution formulation. For an overview of participants and presentations see [1].

This report strives to give an overview of both what happened during the week, and what the outcome of the seminar is.

2 Who is an Insider?

One of the most urgent quests of the seminar was to try to identify the characteristic features of an insider. To this end two research groups presented their taxonomies for identifying insider threat, or insiders per se, and several recent cases of insider actions were discussed—Binney vs. Banner, a message flood created as consequence of a security bulletin, spies that stole secrets for the Chinese Army, and a tax authority employee who used her influence to embed backdoors into taxation software (see boxes below for short summaries). While these cases could not differ more, they served the purpose of illustrating the widely differing characteristics of insider threats, and initiated an intense discussion of how one possibly could aim at detecting and inhibiting them.

The wide range of properties that can characterize insider threats was confirmed in the discussion of taxonomies in Section 3. Especially Case 2 (message flood) caused a lot of discussion, as it seemed unclear whether this really was an insider case, and if yes, whether this was the deed of a single insider, or a confluence of several actions by insiders. Case 4 (taxation software), on the other hand, seemed typical for an employee who is an insider, but needs to "break into" the system to reach certain goals. As result of the discussion we identified the need for further work on: 1) a common vision of how insider can be categorized; and 2) security policies for countering insider threat, and on how to evaluate the impact of alternative security policies.

In the discussion, several definitions of what characterizes an insider were suggested

- An insider is defined with respect to a resource, leading to "degrees of insiderness";
- An insider is somebody with legitimate access to resources;
- An insider is a wholly or partially trusted subject;
- An insider is an individual who has or had access to resources;

Example 1: The Hard Disk Example: Naive user and absent policy
On April 5, 2003, Banner Therapy (a small privately owned company in North Carolina, USA) employee Christina Binney was discharged from her position for "misconduct", and instructed not to return to the office.
Christina Binney was also a co-founder of Banner Therapy. According to Banner, there were two reasons for Binney's dismissal. First, the company disputed her assertion of copyright interest in the company catalogue and website. Second, the company claimed she impermissibly removed from her work computer a hard drive that she took home over the weekend to prepare for a client meeting. The company claimed that the disk drive removal crippled Banner's operations and placed vital company data at risk. Binney explained that a Banner customer requested a meeting on a Friday for the following Monday morning. To prepare for the Monday meeting, Binney chose to physically remove the entire hard drive from her work computer to use with her compatible home computer, rather than take time to transfer the files to a disk.
At the time, Banner Therapy had neither company policy about taking work equipment home nor established computing protocols. When Binney attempted to return to work on Monday, she was denied access; this inability to enter the workplace prevented her from returning the hard drive as she claimed she intended to do.

Example 2: The Email Example: Ordinary user generates an extraordinary amount of email
In early October 2007, Alex Greene was changing jobs. In preparation for the switch, he wanted to update his subscription to a Department of Homeland Security intelligence bulletin by changing his designated email address. In doing so, he mistakenly hit "reply all", and touched off a listserv free-for-all when his request arrived in the electronic mailboxes of several thousand government and private sector security specialists. The result was what commentators described as a mini distributed denial of service attack. There were more than 2.2 million emails pinging among approximately 7,500 recipients before the email server was forced to shut down.
The information contained in the bulletin is unclassified, but nevertheless, the decision to respond inadvertently compromised classified contact and departmental information. Individual subscribers with security classifications remained anonymous until they also hit reply, responding from work accounts that included automatically generated signatures. Indeed, one poster pointed out that, armed with the information contained in auto-signatures, he was one fake letterhead away from impersonating a Department of Defense employee.

Example 3: The Trade Secret Example: Malicious user steals trade secrets
On June 16, 2007, FBI agents, using a sealed grand jury indictment, entered two luxury homes in Silicon Valley and arrested a pair of engineers. Both Lan Lee (an American citizen) and Yuefei Ge (a Chinese national) had worked for NetLogic Microsystems (NLM) until July 2003. The two men used money from mainland China to create and incorporate a company for the sole purpose of exploiting the secrets they stole.
Lee and Ge downloaded sensitive NLM documents onto their home computers. NLM data sheets are "top-level confidential technical descriptions of their products", including information described in enough specificity to enable someone to produce the technology. Together, the men accumulated the information needed to design and produce their own lines of microprocessors and microchips. To finance the business they were creating, the men contacted Beijing FBNI Electronic Technology Development Company Ltd, and entered into an agreement to develop and sell microprocessor chips. Both men were able to access proprietary information without exceeding their individual authorizations.
By late September investigators had uncovered evidence that the venture capitalist had ties to the Chinese government and military.

Example 4: The Tax Fraud Example: Perimeter definition and system design

The District of Columbia (as of summer 2008) is pursuing a case against Harriette Walters and her co-conspirators, for perpetrating the biggest fraud in the city's history. Until her arrest, "Walters was a 26-year tax employee known among her colleagues as a problem solver with a knack for finding solutions by using the department's antiquated and balky computers or finding a way around them." She allegedly used her position to produce fake checks for bogus refunds with fictitious names; the total is said to exceed (USD) $50 million.

The scheme involved Washington's new Integrated Tax System. During design phase, Walters "contributed to the decision that her unit, which handled real estate tax refunds, be left out of it." At the time, the decision seemed to make sense. D.C. had spent $100 million to implement the business and income parts of the system, and it had only $5 million remaining for implementing the real estate tax portion. So the system's perimeter was defined to omit real estate tax processing.

That design decision allowed Walters and her co-conspirators to create bogus tax refunds with fictitious names that were not checked against actual real estate records. Some refunds were issued multiple times; the recipient (often someone's boyfriend) would claim that the check was never received, and a new one was issued—with interest to compensate for the long delay! The schemes exploited several loopholes: each check was under the $40,000 threshold for requiring a supervisor's approval, and no action was taken to cancel the first check or confirm that it had not already been cashed.

- An insider is a system user who can misuse privileges;
- An insider is an individual with authorized access who might attempt unauthorized removal or sabotage of critical assets or who could aid outsiders in doing so; and
- An insider is a person or company whom we trust.

These definitions led to a series of discussions on what we mean by "access" (code, credentials, timing of access rights), whether an insider is sufficiently defined based on resources or whether a definition should take the system into account, and how the definition relates to a masquerader, namely an outsider being able to trick a system into believing he is an insider.

Having these aspects in place enables us to reason about what makes a good insider:

- Knowledge, intent, motivation
- Possesses power to act as agent of the business
- Knowledge of underlying business IT platforms
- Knowledge/control over IT security controls
- Ability to incur liability
- In pecuniary terms
- In brand damage or other intangible terms

The skill of insiders was mentioned, but not thoroughly explored, as a factor 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. A key observation is that the definition of an insider for threat purposes is different than the definition for business purposes.

To summarize, discussion centered on the question of whether an insider is defined in terms of someone with:

- Knowledge: Implies an open system, one that remains secure (if at all) even with full knowledge of the system operation; alternatively, security through obscurity; or
- Trust: An individual is empowered by the organization to be an insider; or
- Access: An insider is in possession of a credential giving access to the system — an IT centric perspective, since the system in general does not know who possesses the credential.

At the end of our seminar a trust based definition of an insider was proposed:

> "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 definition 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 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.

"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.

Insider Threats. A natural consequence of having defined the term "insider" is to consider the term "insider threat". As to be expected, again several aspects were suggested:

- Risk to organization or organizational resources posed by actions (the entity as agent of the business)
- Access type or system role
- Aim or intentionality or reason for misuse

- Level of technical expertise
- Type of malicious insider behaviour or system consequences.

Alternatively, a different framework for insider threats was suggested: there are masqueraders (individuals pretending to be legitimate insiders but without valid access); traitors (legitimate insiders acting in malicious ways) and naive insiders (who cause damage without malicious intent). This is very closely related to the question of motivation discussed above. The problem of interest is dealing with the "real real insider" — an individual deeply imbedded in an organization (e.g., a high level executive, or a systems administrator). Detection techniques for each of these types of insider threats will vary.

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. An important aspect here is that "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.

3 Taxonomies

A key conclusion of the seminar, and actually one of its initial motivations, was that an effective taxonomy of insider threats is a necessary foundation for further work by both researchers and practitioners. Taxonomies provide a means to order problems; such ordering is necessary both to differentiate types of insiders and types of insider threats, and to make explicit the key dimensions which serve as the basis of the differentiation. By identifying the key dimensions, we can then begin to systematically build prevention/response strategies. At least some common acceptance of the dimensions of a taxonomy is needed; experts can disagree about how the dimensions are applied (as in defining who is an insider) but by forcing an explicit discussion of different interpretations, taxonomies can serve a vital role. Consequently, much of the discussion of taxonomies overlapped with the discussion of who is an insider and how do we define insider threats.

Two different formal taxonomies were proposed. The first proposed taxonomy observed that insider threats can be defined in terms of four reference perspectives, namely the organization, the individual, the information system, and the environment. This "holistic" or top-down taxonomy was then used to illustrate how it differentiates the four insider threat cases described before. The second taxonomy defined 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, a concept picked up several times throughout the week.

Key factors important as determinants of the insider threat may be difficult to categorize a priori. As noted elsewhere, knowledge of the insider's intent is desirable, but requires all-embracing knowledge. Several less formal taxonomies were proposed during the seminar. It was suggested that the distinction between malicious and accidental threats is misleading, and that the real distinction was between doing something intentionally (for malice, or good reasons which nonetheless may result in damage) versus events that occur accidentally. Along with intentional versus accidental events, a second distinguishing factor proposed was whether the act was obvious or stealthy. An alternative view classed insider threats into masqueraders (those possessing proper credentials illegitimately obtained), traitors (insiders with legitimate credentials who maliciously misuse their access) and naïve or accidental use that results in harm. A third perspective proposed would categorize insiders along four dimensions: access type; aim or intentionality or reason for misuse; level of technical expertise; and the system consequences of insider threats.

The level of knowledge or skill that an insider has shapes the frontier of threats that can be posed by that individual. Knowledge levels can be classed in increasing order as that possessed by outsiders, ordinary insiders, and privileged insiders. It was noted that masqueraders, lacking the knowledge of insiders, may be detected by the inordinate searching required because of their relative lack of knowledge of organization resources.

Notable in retrospect by its absence was any discussion of how or indeed whether a single taxonomy should be adopted by the community. In some cases, like the taxonomy for speciation (kingdom, phylum...) a more or less universal adoption has provided great value. This question will be considered for inclusion in future seminars.

4 Issues in Detection/Forensics/Mitigation

Forensics appears to be highly undeveloped when addressing insider threats. It was noted that we have no clear idea of forensic data needed. Insider behaviour may be close to the expected behaviour, and the still often used audit trail is generally inadequate (redundant, misleading, missing data)—even worse, usually audit trails lack time correlation. The number of appropriate characteristics to observe may be large: in other circumstances like credit cards up to thirty-five characteristics are used to establish legitimate behaviour . 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”. 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.

Detection, forensics, and response must also wrestle with how to distinguish motive and intent. Malicious acts may be equivalent to acts due to accidents or naïveté. Insiders may legitimately use domains in unexpected ways that might trigger false alarms. Outsiders, insiders acting with malicious intent, insiders acting without malicious intent, and accidental behaviour may all result in similar effects on the organization. Hence there are always going to be gray areas in how security policies define both insider misuse and proper behaviour. Furthermore, actions are context bound, but most security polices only inadequately capture the nuances of context.

Monitoring. While monitoring can help with technical aspects, it does potentially worsen behavioural aspects. The deciding factor is how much monitoring is acceptable (both ethically and legally), and is it at all beneficial. The noteworthy point here is that this question arises at all levels from individual actors, to groups, to companies, to the society as a whole. The problem is that not only may the same actor have different opinions depending on at which level he is asked, but also that different answers for different individuals may exist at the same level.

One should observe that 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 [2,3] had in common that the attacker intimately knew the monitoring system and knew how to play it. However, it was questioned whether in fact malicious insiders seek to avoid setting off monitoring alarms by slowly adjusting their profiles, as is widely hypothesized.

In summary trust into insiders is a behavioural expectation that, so was the consensus, still needs to be controlled. While the easy solution to reducing the number of insider cases would be to remove all restrictions (making the illegal actions legal by means of removing the semantics of the term “legal”), we aim for making the monitoring as efficient as possible, where in different situations the term “efficient” may have different interpretations.

The goal of monitoring (or observing in general) should be to only monitor what is needed to identify the threat in question. Since currently trust can often be transferred, for example by handing over a code card, it is important to isolate transferred trust as much as possible, not least to allow the result of monitoring to be used to bind actions to actors.

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. Observations on “higher level behaviour” 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 [4] that in most cases someone else knew about the insider threat actions going on, so it is important to find ways to encourage reporting of “suspicious activity” by others? Looking for suspicious (different) behaviour patterns by insiders is appealing, but difficult to systematically apply; behavioural 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 behavioural changes is bound to produce false positives from otherwise innocent factors like individual predispositions or lifestyle changes. Fundamentally, the attributes key to insider threat identification will be largely context bound.

From a company point-of-view it turns out to be often preferable to not mitigate ongoing insider attacks for numerous reasons. Here optimistic access control is seen as a viable option, i.e., allowing insider actions 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. It seems often more ruinous to take systems down because of an ongoing attack than to accept the losses and prosecute after the fact.

Outsourcing. Even more difficult is the handling of outsourced parts of a company, both for technical and legal reasons. On the one hand data may be much harder to obtain (or be less trustworthy), at the other hand the data protection laws regulating data collection in other countries may be vastly different from the laws in the home country—it should be noted that this might actually beneficial. Collaborative fraud detection was mentioned as one of the counter measures to be applied in such a case.

Another problematic area is outsourcing the auditing itself; while in the "regular" outsourcing scenario it may be difficult to obtain the data in the first place, it now is paramount to protect the already collected data. This means the data should be anonymized as much as possible, revealing only as much data as necessary to allow external auditors to produce meaningful results, but at the same time hindering them from drawing meaningful information from them. One example is to anonymize timestamps to still reveal relative ordering between events, but blurring them such that the exact order and timing is lost. It was noted that formal methods can and should be applied in these settings, and some were presented, but at the same time often require significant resources, such that for example before and after the fact application is often feasible, but not online detection.

Coming back to the insider classification into masqueraders, traitors, and "regular" users, one can now describe the required detection techniques—for masqueraders it seems most important to perform profiling and anomaly detection, and for traitors a trap-based decoy network was suggested, identifying access to fake documents. As pointed out above it is of utter importance to avoid false positives when trying to identify masqueraders, who can be expected to lack knowledge about real persons' behaviour and there probably will perform inordinate amounts of searching.

Comparing the insider threat in companies with that in public agencies or the military it was noted that there are significant differences, especially due to special employment conditions. 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. This is often (and must be) supported and accompanied by a legal framework that is unique to these settings and aware of the special circumstances. This setting opens up completely new problems, that were not discussed in full detail.

5 Policies

Policies obviously play an important role with respect to insider threats, as they define the boundaries between permissible and not permissible behaviour, both on a technical and non-technical level. They not only define proper behaviour, but implicitly also define the notion of insider. It is problematic that policies often are only specified implicitly, possibly leading to large differences between a defacto policies and a "real", intended policy. To support the externalization of these intended policies, policy languages have been developed, which are usually quite well suited to support technical issues, and at the same time try to add support for non-technical aspects of policies.

During discussions the issues of "context" and "dynamicity" came up several times. For example, a given actor might be an insider in one situation, but would be considered an outsider in another. 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. This ties policies as well as their specification and enforcement tightly to human factors, which are discussed in the next section.

Policy Hierarchy. Policies themselves are developed based on three sources: 1) legal and regulatory (so-called best practices); 2) business requirements; 3) security requirements. It was noted that all of these sources can result in implicit or explicit policies, establishing a grey zone where behaviour is neither good nor bad. This potential gap is extended and formalized in the Unifying Policy Hierarchy [5], 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.

Gaps and conflicts in policies came up repeatedly in the discussions as 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.

To illustrate this the hierarchy was applied to the example Case 4 (DC tax system). At the oracle and feasible level the design contained auditing throughout the system, but at the configuration level this was dropped due to budget cuts. Similarly, at the oracle level refund checks should only be issued in the case of lost checks, but due to constraints at the feasible level, refund checks are issued on request.

Another example is attraction of interest by applying security classification, 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 analyse 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 behaviour to be permissive.

Policy languages. A gap of a different kind exists for policy languages. Here the gap exists between the existing capabilities to specify system (and more broadly, organizational) policies, and the needed qualities of policy to adequately prevent insider threats. One of the most urgent needs, already mentioned above, is for policies to be aware of behavioural aspects and context, thereby being able to handle and regulate abstract events. While a "zoo" of policy languages exists, with a vast overlap in terms of what they can achieve, the user often may not be able to write policies, let alone read, understand, and follow them. As an example it was mentioned that insiders in XACML using delegation might actually be able to perform denial-of-service attacks on a policy engine, be it by innocent misuse, or be it to block a policy resolution service. This is expected to gain growing importance in future interactions between previously independent parties.

As mentioned above many systems define the notion of insider relatively to system boundaries. In the long run we may therefore need domain-specific policy languages, in which for example actions would be allowed only if discretionary circumstances justify their execution.

From a research perspective we 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 for compliance with security practices for the sake of security. However, having some policy is not enough, that is deploying security technology and policies does not automatically help you in achieving security. This is especially true due to the above-mentioned context-dependency of policy rules.

Structured Exercise. As part of the seminar we played a structured game, in which three teams each in rotating roles first developed a threat scenario, then passed their threat scenario off to a second team who designed security policies to address this threat, and finally to the third team to evaluate (and possibly find defects in!) the proposed policy response. From this exercise the participants came to three conclusions: first, that technical approaches devoid of or mismatched with underlying business and economic principles of the organization the security did not work; second, that it is hard to determine if a solution will work without a budget (a consideration that was little discussed during the seminar); and third, that the game illustrated the difficulty of quantifying the risk exposure facing an organization from an insider threat.

As a result of these discussions we assembled a list of research questions related to policies:

- How to keep statutory best practices in line with real best practices;
- Risk analysis for bad legislation/regulation;
- Risk analysis when the magnitude of the risk is uncertain/unknown/unknowable

It was also noted that there is very little work on bridging the gap in policy levels and understanding.

6 Human Factors and Compliance

Three foundational observations shaped the discussion about human factors and compliance. First, it was noted that most insider threat policies are based on a set of incorrect assumptions, namely, that 1) once someone is vetted, audit and incident management processes will pick up policy violations; 2) that risk assessment processes will pick up changes to individual and group values; and 3)that training and awareness-building education programs will instil desired security culture. However these assumptions never fit to how people actually want to pick up information, or act in the course of doing their jobs within the organization. The inadequacy of many existing security solutions to address real life human behaviour presents us with a set of challenges on how to better incorporate human factors into solutions.

Second, a general theme in many presentations and discussions was the need to align security policies with organization workflow, or, stated simply, security should support people doing their jobs. Several examples were cited of technological security approaches that, because they interfered with work flow, were not accepted and in fact actively subverted (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 you get to the limit of peoples' tolerance for disruptions to their work flow and social interactions. There was considerable discussion about how to best incorporate usability into the development of security policies and application of security technologies. These issues have been addressed in other applications (outside of cyber security) so that one conclusion reached was that insider threat research should look to other fields for experience in addressing these challenges. Finally successful security policy needs to demonstrate to insiders the value of security, not just the requirement for security.

A third observation was that motive and intent matter a great deal, but multiple motivations may map into a single intent. One example cited was 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 but not the motivation. This has important implications on the limitations of monitoring, and highlights again the need to establish context for specific actions.

Understanding and Integrating Human Factors. 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 behaviour: for a crime to be committed a person must have both motive and opportunity. As just noted, motive matters; for the "cold intellectual attacker" when the possibility of punishment is high and the sanction severe potential criminals will be deterred from committing illegal acts, especially when their motives are weak. More generally, 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.

Organizational purpose and management structures affect both security structure and policy. In discussing organizational factors relevant to the insider threat a number of questions were raised:

- 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 are role for technology in extending the whistle blowing capabilities of staff?
- Policy conflict within organizations: The discussions concluded 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? To do so might require a hierarchy of organizational needs like the Maslow hierarchy of individual needs.
- 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 behaviours?

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

Insiders (or people in general) will act unexpectedly. One participant noted that , "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 behaviour.

Hence, 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.

The unpredictability of human behaviour has its implications for the role of trust in an organization. Trust is a behavioural expectation, and trust is only necessary in an organization when behaviours cannot be controlled in all dimensions. Trust is also transitive. Some participants argued 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.

Unpredictability may also be an artifact 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 won't put up with monitoring."

Policies and Human Factors. Policies need to be shaped and evaluated in terms of their human impact. How specific should policies be? 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 behaviours, and fit with business processes, and organizational values and norms. A common theme is the need to link the user community (the insiders in the organization) with the policies being developed and enforced. Failing to engage staff in security may be the norm (though this observation, implicit in many comments, was never explicitly discussed) but this lack of engagement weakens security. Security will only work in organizations where people feel that they are part of a larger community. One suggestion was that organizations should conduct specialized internal exercises 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.

Issues of what we can measure, what is relevant to measure, and how and when we intervene when suspecting threatening insider actions need to take human factors into account. 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 behaviour, 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!), or be affected psychologically, the organization’s culture may be affected, possibly for extended periods. There is, therefore, a need for decision processes to decide when to intervene and how.

7 Surveys and the Real World

A recurring question was whether researchers and practitioners have sufficient data to understand the range of insider threats, formulate appropriate prevention/detection/mitigation strategies, and subsequently evaluate their effectiveness. The problem, of course, is that most organizations deem insider threats too sensitive a topic to discuss it openly. While many seminar participants perceived the lack of data to be a significant roadblock to further progress, some participants did not, suggesting that we already understood much of the nature of the problem, and that informal feedback from practitioners will suffice.

Data from ChoicePoint was referenced to indicate the ubiquity of insider attacks. While fraud is perceived as something that only happens to others, companies experienced annually two to ten cases of insider abuse (average is 5), with an average loss per company of (USD) $159,000. However ubiquitous insider threats are, the data we have available is poor, and not just because of the reluctance of organizations to share their histories. We look only at people who get caught, and have little real understanding (other than anecdotes) about the baseline population of threats. How applicable can these studies be? Furthermore, we have little insight into the baseline population of all insiders, making characterizations of threatening insiders (who are caught and reported) meaningless if we cannot reliably differentiate between the two populations.

Two different approaches to developing data were presented. One family of approaches relies on decoys or simulations to capture real, or quasi-real insider threat behaviour. A decoy document (or trap-based decoy) infrastructure using honeypots and honeynets already is being deployed by one participant to generate data on insider attacks. For traitors (malicious insiders with legitimate access) this approach was outlined as being useful. Decoy documents have to be structured carefully, so that they are differentiable by the user/owner as fake, are non-interfering with normal operations, and "variable" in nature and constantly changing when generated. However, masquerader attack detection is hard, and generating real data on these types of attacks is hard, too. Running simulations with subjects instructed to act as insiders in certain ways (e.g., malicious or non-malicious) is also being used to generate data, but obviously sub-optimal.

Another approach modeled systems at a generalized level. Generalizing from the current approaches which depend on logging actions and then performing a positive audit of log files to check for "conspicuous events," the example system is modeled (in tuple space) by locations and connections to form a directed graph. Domains are defined upon which restrictions and capabilities are imposed to model access control, and then actions are performed on the system. While not, unlike the first family of approaches, generating real or quasi-real data, this modeling helps determine what observables can be obtained at all stages of the insider threat process.

A second set of issues revolves around our ability to predict insider threats. As noted, attempts to base prediction on behavioural or other personal characteristics suffer from small sample sizes and the inability so far to adequately distinguish insider threat characteristics from those of the baseline population. In general, statistical methods for predicting insider attacks face a number of challenges: probabilistic methods demonstrate correlation, not causation, and there is always the problem of analyzing (rare) cases or unusual behaviour. Consequently, most alerts based on predictive modelling are false positives. The reaction strategies by decision makers to alerts (ignore when too many? Always escalate to further investigation in order to "cover their backs"?) may vary in their effectiveness, but this issue was not pursued during the seminar.

Overall conclusions about the use of predictive statistical models were 1) take care to get the statistics right (by using control groups and other careful methodology) and 2) be mindful of the analog to the Heisenberg principle, that is, by monitoring (and investigating) people, behaviours change.

A persistent question about the "real world" of insider threats is whether this problem is uniquely American. Certainly most attention and research on insider threats has been American. Some seminar participants concluded that maybe yes—if effectively addressing insider threats is inherently context dependant, then the American concept of the generic manager (exemplified by GE's management style) lends itself to having more insider threats. If insider risk depends on context, then we need more than generic risk managers to effectively address the problem.

8 Conclusion

The Dagstuhl seminar on Countering Insider Threats provided a week-long base for discussions of what the relevant aspects of the problem are for different communities. In motivating the seminar, we asked the seminar participants to jointly emerge with a definition of what or who an insider is, and how to deal with the threat posed by it. It is our believe that we succeeded in getting a better understanding of what these different communities mean by “insider”. As stated above, this knowledge has already during the seminar been used to develop integrated approaches towards qualitative reasoning about the threat and possible attacks. Beyond this shared definition of what constitutes an insider, the most prominent outcome of the seminar is the beginning of a taxonomy or framework for categorising different threats. While this fell short of the ambitious goals we originally had formulated [webpage], the process of reaching this definition was highly enlightening and is documented in this article. In the process, the seminar identified the need for:

- A framework or taxonomy for distinguishing among different types of insider threats;
- Methodologies for assessing the risk and impact of insider threat incidents;
- Incorporating human factors into the development of solutions;
- Better formulations for specifying useful policy at both systems and organizational levels—policy that would be meaningful and applicable to the insider threats deemed most important.

There were some cross-cutting conclusions that emerged from the seminar. The role of trust was discussed in a number of different contexts. In one sense, the ideal security framework for addressing insider threats would eliminate the need for trust — all behaviours would either be defined permissible, or else made impossible to execute. But this model ignores two realities. In any but the simplest settings, context of actions is highly determinative in shaping what is appropriate or needed behaviour. Further, many (most?) organizations would not accept a working environment so rigidly defined as to eliminate the need for trust. Hence, we emerge with the conclusion that trust relationships will be present in most organizations; how to best factor trust into security policies and frameworks remains, however, unclear.

Security, moreover, is context dependent. Security is not achieved by deploying generic (context free) controls. However, the importance of context in addressing insider threats poses a number of challenges. Capturing qualitatively the various situations that might arise in an organization is itself probably impossible, though effective dialogue between those defining security controls and those working as insiders in the organization will certainly help. Hence, insider threat prevention and response has to deal with the reality that controls will not adequately capture all of the behaviours that might be appropriate in a given context. Even if all contexts could be qualitatively described, policy languages and controls are inadequate at the current time to fully capture the range of contexts identified.

Motivation and intent clearly are important in defining insider threats and defining appropriate detection/forensics/mitigation strategies. While intent (the purpose of actions) is at least partially observable, motivation (the incitement to action) is not. The intent to, for instance, obtain certain data may reflect malicious motives, or may reflect positive motives (as in a hospital emergency where certain information is desperately needed regardless of legitimate access).

Devining motivation highlights the need for context aware policies, but even with context motivations may be difficult to determine. We conclude that approaches for understanding motivation a priori are still highly immature.

Each of these observations emphasize the conclusion that security will not be achieved solely by deploying security technology. Most people are not entirely logical or consistent in their behaviour, and this confounds our ability to formulate measures to reliably prevent or detect malicious insider behaviour.

As we write this report we are in the middle of preparing an application for a follow-up seminar We would like to thank all participants of the seminar for making it a fruitful and inspiring event—and especially Dagstuhl's wonderful staff, for their endless efforts, both before and during the seminar, to make the stay in Dagstuhl as successful as possible.

8.1 Community Development

As stated above we believe that the week in Dagstuhl has been influential in heightening awareness among communities for activities and developments. During the seminar many participants expressed the wish for a community website to establish a central focal point, both for communication between communities, but also to the outside, governmental agencies, and companies. This web portal is currently under construction [6].

References

1. Homepage of Dagstuhl Seminar 08302: "Countering Insider Threats". Available from http://www.dagstuhl.de/08302, last visited December 4, 2008 (2008)
2. Wikipedia: Jérôme Kerviel. Available from http://en.wikipedia.org/wiki/Jerome_Kerviel, last visited December 4, 2008 (2008)
3. Wikipedia: 2008 Liechtenstein tax affair. Available from http://en.wikipedia.org/wiki/2008_Liechtenstein_tax_affair, last visited December 4, 2008 (2008)
4. Cappelli, D.M., Moore, A.P., Shaw, E.D.: A Risk Mitigation Model: Lessons Learned From Actual Insider Sabotage. In: Computer Security Institute, 33rd Annual Computer Security Conference and Exhibition. (2006)
5. Carlson, A.: The unifying policy hierarchy model. Master's thesis, Department of Computer Science, University of California, Davis (2006)
6. Countering Insider Threats: Community website. http://www.insiderthreat.org (2008)