The Risk of Risk Analysis
And its relation to the Economics of Insider Threats

Jeffrey Hunker
Carnegie Mellon University
jhunker@andrew.cmu.edu

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

Abstract

Insider threats to organisational information security are widely viewed as an important concern, but little is understood as to the pattern of their occurrence. We outline an argument for explaining what originally surprised us: that many practitioners report that their organisations take basic steps to prevent insider attacks, but do not attempt to address more serious attacks. We suggest that an understanding of the true cost of additional policies to control insider threats, and the dynamic nature of potential insider threats together help explain why this observed behaviour is economically rational. This conclusion also suggests that further work needs to be done to understand how better to change underlying motivations of insiders, rather than simply focus on controlling and monitoring their behaviour.

1. Introduction

The *insider threat* or *insider problem* has received considerable attention, and is cited as the most serious security problem in many studies.[1] It is also considered the most difficult problem to deal with, because an insider has information and capabilities not known to other, external attackers. Examples for insider threats are manifold (*e.g.*, [2, 8, 10]), but usually only those resulting in significant harm are noticed by the public.[2]

For example, in January 2008 Societé Generale suffered a $7 billion equities derivative loss due to the activities of a trader who had moved from the back office of the bank to become an apprentice trader in the dealing room. As the newspaper Liberation noted,"the stars of finance must be very cross that a simple base trader has succeeded in sinking a bank. The fraud is terrible for the credibility of the bank in the equities derivative sector, a business in which Societé Generale has become a global leader" [13, 17]. Of course only vague details of the case were revealed, and we will probably never know the exact details of the case, but the little we know hints at insider actions being responsible for the considerable damage.

Or for example Christina Binney, a senior employee of a small company, Banner Therapy, who without violating a specific company policy took home for the weekend the company's hard drive. She was subsequently fired for this action, the company claiming that her action put the company's very existence at jeopardy [5].

In a third highly public example, the US District of Columbia is pursuing a fraud case against a middle manager who used her influence to exclude her unit, dealing with real estate tax refunds, from a new Integrated Tax System. This exclusion allowed her to create bogus tax records that were not checked against actual real estate records [14].

In contrast to these high profile cases, lesser damages caused by insiders usually are covered up even if discovered. This goes in line with reports by professionals in organisations concerned about insider threats: their organisation is aware of insider threats but takes only limited steps to prevent them, including threats posing the most serious impact. After the event, however, it is often considered crucial to have sufficient proof and documentation to be able to deal with these cases [20].

In the light of such severe consequences one should expect that preventing these threats would be one of the topmost priority for organisations. However, as many senior managers state, their organisation is aware of the threat, but does little to prevent it.

We find this observation to be surprising, to say the least, and in case it is true, which recent events like the ones mentioned above indicate, the question is why organisations choose to be so vulnerable? The answer would be simple if the vulnerability were a matter of sloppiness by the organisation. However, it seams that what we are talking about reflects what is presented by senior managers as a distinct choice.

In this way insider threats fundamentally differ from external threats. Organisations rarely choose to leave open vulnerabilities in their systems that might be exploited by outsiders to destroy or significantly damage the organisation. If organisations do leave open such vulnerabilities, the reason is either limited resources (in which case one needs to examine the substance of the organisation's risk analysis), or sloppiness.

In this paper we discuss the question why organisations, given the importance of insider threats, choose policies that allow insider threats to occur even in the face of adequate resources? Is this decision based on the sense that organisations (or their security personnel) figuratively throw up their hands in the face of a threat that, while recognised, seems impossible to adequately address? Little public data exist to help answer the question of whether such behaviour is economically or organisationally rational.

We develop an answer to this question by examining the relation between an organisation's risk analysis, the assessment of trust in an insider, and how both of them (should) develop over time. We argue that as insiders over time gain more knowledge and thereby become a bigger risk, the organisation only has two choices how to react. Either, the organisation chooses implicitly or explicitly to have more trust in insiders as they pose a potentially bigger risk, or the organisation needs to apply and enforce an ever-increasing number of policies to regulate the insider's actions.

In the rest of this paper we lay out a series of observations (based on anecdotes and extensive consultations with both researchers and practitioners), from which we derive a framework for understanding this observed behaviour, and its implications for strategies for

[1] For example, in a 2007 Computer Security Institute survey about computer crime and security, 59 percent of respondents perceived that they had experienced insider abuse of network resources [9].

[2] Since 1995 only 119 cases of insider threats prosecuted under US Federal law have been identified [11].

dealing with insider threats. We develop a combined view of the economics of the different components in this framework—the organisation, the insider, and elements of mitigation all have a combination of goal function, risk function, and/or cost function associated with them. This obviously results in a multi-dimensional optimisation problem, whose complexity eventually explains that our standard tool for assessing threats, risk analysis, breaks down in the face of one of the most vicious threats.

Our main conclusion will be that "complex" insider threats emerge as insiders with malicious intentions adapt their behaviour to circumvent control systems. They often succeed because they have intimate knowledge of the control system, and especially of its blind spots. This adaptation of behaviour makes it almost impossible to detect and prevent insider threats, and leads to a high uncertainty about possible threats, and in turn renders preemptive actions prohibitively costly. This benefit/cost ratio ultimately is the reason for organisations to refrain from defending this kind of insider threats. We discuss some possible measures how to prevent these complex threats from occurring. For a discussion of risk and uncertainty see Knight's seminal work [16].

While most of this paper considers insider threats, many of our results are applicable just as well in any risk/threat scenario, which involves trust. For a discussion of models for explaining insider threats see, for example, [22, 21].

2. Insiders, Outsiders, and their Threats

While there is no commonly accepted definition of either an "insider" or an "insider threat" recent work [12, 20] points to a trust-based definition of an insider:

> *"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 organisation's structure."*

The rationale behind this definition is that it removes any specific IT bias from the definition, and focuses on organisational assets rather than a narrow approach based on system credentials. The insider has been legitimately empowered to do some things that affect the organisation, and he is trusted to use this empowerment wisely in a way that will benefit the organisation, or at least not harm it. Beyond this definition [20] identifies factors of a "good" insider:

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

All of these are affected both by time and position within the organisation.

As mentioned in the introduction, insiders obviously have a special role for an organisation. While an organisation in general will try to do whatever possible to prevent threats from the outside, it often can or will not do so with threats on the inside. In the next section we present a series of observations, clarifying the relation between trust and risk, and their role for internal threats.

In contrast to insiders, outsiders usually are easily identified, as is the amount of access they should have to an organisation's data and assets. The clear separation of concerns between outsiders and an organisation eases controlling interactions with outsiders by means of access control and policies. It should be noted that above definition of insiders elegantly solves the problem of outsiders having special rights on an organisation's assets—since they have been granted access, they are correctly treated as insiders.

Before further investigating the role of an insider in an organisation, we first define what we mean by "insider threats". Insider threats emanate from individuals who are insiders according to our definition, and whose actions place the organisation at risk. These actions can be maliciously motivated, the result of accident or error, or made because the individual is deceived. The insider threat can be caused by an insider acting alone, or in concert with other insiders, outsiders, or various combinations of the two.

Thus, insider threats encompass a wide variety of different types of actions that can have a correspondingly wide range of impacts on the organisation. While work on developing complete taxonomies of different insider threats is underway [6, 12], a simple categorisation, sufficient for our purposes, is to differentiate by motive and complexity of trust relationship:

- For *motivation* we distinguish between accidental and intentional actions; and
- for *complexity of trust relationship* we distinguish between simple and complex ones.

Based on this categorisation, we consider the following scenarios of insider threats.

2.1 Insider threats that do not represent a violation of trust:

- *Accidents or stupidity*: People will be stupid and we cannot anticipate stupidity or accidents very easily. There is considerable work on ways of anticipating and preventing such instances, and much of it draws on work in fields (like nuclear plant operation for example) where the lessons are nonetheless applicable to the insider threat issue [18].
- *Fulfilment of duty*: Organisation's policies tend to get in the way of performing a task. Insiders may decide to disobey a policy, and thereby on the one hand be able to fulfil their duty, on the other possibly causing an insider threat, which they might or might not be aware of. With Binney and Banner Therapy, Binney apparently was unaware that she was potentially threatening the organisation's survival. Considering the trust-based definition of insiders given above, they are trusted to judge whether or not the situation justifies breaking the rules [24, 1].

2.2 Insider threats that do represent a violation of trust:

- **"Simple" insider threat**: The typical example for this is the disgruntled employee, who might be at risk of being fired and causes damage to the system, or steals some files they have access to; or a person who is paid to steal data. Most of these are cases where the system facilitates the damage, *i.e.*, the same damage could have been caused in a pen and paper system, or the threat involves violation of trust that is not easily picked up on, *e.g.*, an employee reading printouts in a printer room they have access to, or the recently fired employee who still has (through administrative oversight) access to the organisation's computers. The key property of these cases is that the damage done to the organisation while potentially considerable, also reflects a violation of a simple trust relationship.

"Simple" insider threats depending on violations of trust can be thought of as follows: the losses caused are not too high, and can therefore justifiably be ignored; or the potential harm is considerable but the threat depends on trust relationships being violated in a simple fashion, meaning that they could easily have been prevented. In either of these cases the organisational response is appropriate—we either absorb the cost as part of doing business, or revise our security policies so as to avoid a repeat of the insider threat again.

- **High profile (or charismatic) insider threat**: This is the type of insider threat that usually is reported on in the press—the one everyone is fascinated by. Examples include the aforementioned French trader [13, 17], the D.C. real estate tax fraud [14], or the Danish case of Stein Bagger [23], who used his position to build up a complex system of fraud and deception.

These high-profile insider threats with devastating consequences can represent extremely clever schemes. What is more, the insiders causing them usually have more information than the typical insider. As the head of the D.C. tax office commented after the real estate tax fraud was discovered, "Our system has got a plethora of internal controls on it. On top of that, we have manual controls. But you're always vulnerable to an enterprising employee who knows how the controls work." [14].

We hypothesise that these intentional malicious insider threats with large impacts on organisations occur more frequently than appears in the public eye, but the risk of their occurring is accepted by the organisation as an unavoidable risk of doing business.

This hypothesis is based on anecdotal evidence from discussions with private sector and government managers, from the public record, and on the observation that the high level of interest in preventing insider threats by many financial institutions suggests that the problem is viewed as being very serious.

In the rest of the paper we argue that charismatic insider threats fundamentally challenge the basis for risk analysis. In its simplest form, risk analysis depends on:

- Policies[3] directed towards a risk (and their costs)
- Losses due to risks, and
- Probabilities of risks taking place.

We conclude that risk analysis focused on blocking or detecting high-level insiders from carrying out their threats is of only limited value. Alternative ways to increase the confidence that the trusted insider does not become a threat depend on human factors (basically keeping insiders happy); the effectiveness of these policies appears to be little understood in the insider-threat literature.

3. Building up Trust and Risk

Trust is a central ingredient of our private and public life, be it as a person or as an organisation [7], whenever we have to consider a risk. In this section we discuss in detail the relation between risk, trust, organisations, and insiders. In doing so we will repeatedly get back to a mock-up insider story, which illustrates the process of an organisation hiring a new employee, and how he thrives and prospers, turning into an insider and eventually representing a serious threat to the organisation. In Figure 1 we plot the relation between time and the risk that the insider poses to the organisation, and the trust relation between the organisation and the insider. Before looking at trust and risk, however, we first lay out the beginning of the example.

Example: Organisation X wants to hire a new employee. They interview a flock of applicants, eventually picking one.

At this point the organisation has a basic understanding of their future employee, but they do not necessarily have a reason to trust him. This moment is marked by point "0" on the time line in Figure 1—the new employee is not hired yet, the company has neither trust into him, nor does he constitute a risk.

[3] We define policies to mean the set of technical, organisational, and behavioural actions or rules that an organisation has created to prevent, control or encourage actions that affect their information systems. Of course, not all policies are necessarily followed in practise [19], a point we will discuss further.

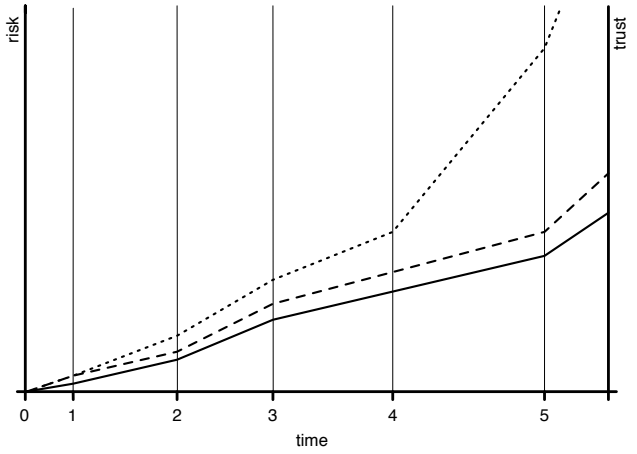

Figure 1. Plot of the **trust** that an organisation has in an insider (solid line) against the risk that the insider poses to the organisation. The dashed line represents the organisation's **acceptable risk**, using tools such as policies and auditing to minimise the distance between risk and trust. The dotted line represents the **effective risk** that the insider emanates. The marks on the time line represent events during the insider's employment in the organisation (see text for discussion). Note that the effective risk could very well be smaller than the acceptable risk.

In order to increase, establish, or justify their initial trust, and assess the potential risk the future employee might pose, they will usually (or at least should) run a background check. Independent of whether or not a background check is performed, once the applicant is hired, the organisation on the one hand establishes a simple trust relation to him, on the other hand he poses a certain risk. Both risk and threat correlate to the position in the organisation he starts at, as well as the assets and data he get access to. This is identified by point "1" on the time line in Figure 1.

For the sake of this section we assume that the insider to be is hired at a rather low-level entrance level. Whatever we describe in the following could just as well occur when joining the organisation at a senior level, which in our classification of insider threats, would represent a complex trust relationship.

3.1 Simple Trust, Low Risk

We now are at point "1" on the time line in Figure 1. The company has established a simple trust relation to the insider, but as just mentioned the new employee also emanates a certain risk for the organisation, part of which might be acceptable.

How does an organisation deal with this situation? To mitigate the risk, and to justify the trust, only simple mechanisms are needed. Based on the established trust, the insider can be granted access to certain parts of the organisation's assets. However, the insider poses a (small) risk to the organisation, and should therefore not be able to freely act in the organisation. A usual mechanism is to control the insider's access to the organisation's assets by means of security clearance, and access rights to certain data and locations. This establishes with help of simple means an easy to control limitation of the risk that the insider can pose.

In this phase the insider's knowledge of the organisation and its assets is fairly limited, and so is the amount of damage he can cause. Over time, this knowledge will increase, and will result in the need to adjust the risk analysis. At the same time, the organisation and the employee develop a hopefully mutual, more complex trust relationship, which to a certain degree justifies accepting more risks.

3.2 Medium Trust, Elevated Risk

Example: After some time the insider changes positions and joins the internal auditing unit, where he works as part of a team that audits the organisation's transactions.

This obviously represents a substantial increase in trust into the employee, and it also means that the employee now represents a significantly higher risk for the organisation, since he gets access to potentially secret data of internal transactions.

On the time line in Figure 1 we are now at point "2"—the trust in the employee has increased, as has the risk that he poses. When considering the complexity of the trust and the risk relation, it has increased considerably, too. This increase is due to the insider's more detailed knowledge about the organisation, both with respect to inner workings and with respect to internal data. As before the organisation may want to limit the difference between risk and trust, by means of a combination of policies, monitoring, and auditing.

The overall situation stays the same as before—the organisation has some trust in the employee, and is willing to accept a certain risk beyond that. As before the organisation may want to limit this risk as well as the potential additional risk posed by the employee, and in this case a typical solution is a set of policies that among others might result in two or more members of the auditing unit being required to access the auditing data, thus spreading the risk over several employees. In contrast to the previous situation, the mix of mitigating factors now is getting more diverse, and potentially more restrictive.

3.3 Complex Trust, even more complex Risk

Example: After having worked in the auditing department for some time, the insider has been promoted again (point "3"), and we meet him some time later, as he joins the trading unit (point "4"), having already established himself in the organisation.

At this point the organisation has built up a fairly high amount of trust into the employee. Due to potentially diverse positions the insider has worked in, and consequently due to potentially manifold knowledge the insider has on internal workings and assets, the trust relationship now is fairly complex. The interplay of different areas of the organisation that the insider has experienced is hard to clearly describe, and even harder to measure.

As a consequence of the trust relationship getting more complex, the risk assessment of the insider will rise in lockstep, as before. However, Figure 1 illustrates that we assume the effective risk to grow significantly larger than the acceptable risk. This is motivated exactly by the fact that the insider has developed a more precise model and knowledge of the organisation, its inner workings, and assets.

Example: While the insider might no longer have direct access to the auditing system, he still knows the details of how the system works and when it is triggered.

For the organisation this can have dramatic consequences. From a "local" viewpoint, whatever policies are applied for employees in the trading unit should work just fine for the insider, since they are tuned to cover exactly the transactions and behaviour that is expected from a member of this unit.

From a more "global" viewpoint, this mitigation of course is completely inadequate, since it does not take into account previous knowledge of the employee. While this problem might be resolvable for transfers inside of the organisation, imagine the effort necessary to identify, assess, and mitigate the risk when hiring somebody from outside into the trading unit.

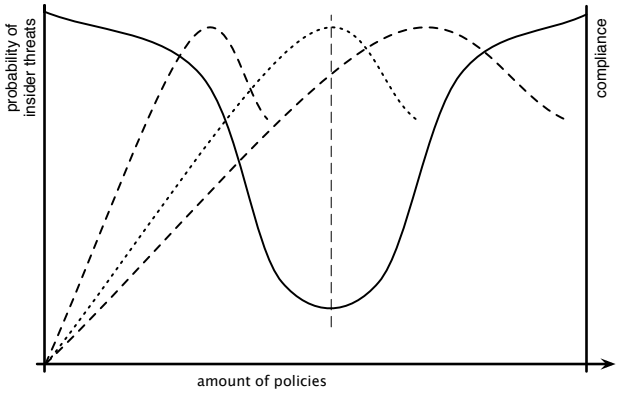

Figure 2. Plot of the number of policies against the likelihood of insider attacks to occurring (solid line), and the likelihood that employees will comply with the policies (dashed/dotted lines) [3, 4]. Organisations want to be at the "sweet spot", where maximum compliance coincides with minimum number of policies (the dotted plot). For compliance, the x axis could also be interpreted as "time passed since a certain policy was introduced", assuming that it takes some time to establish the policy's efficiency, which will eventually degrade again.

Simple trust relationships are relatively straightforward in the ability to control or monitor the risk in our interactions; more complex trust relationships on the other hand pose difficult problems in terms of how to ensure or monitor some degree of trust. While this kind of trust relationships pervade our whole existence, it largely depends on situational factors how much we rely on them in making decisions.

In any kind of relationships we therefore face a number of problems related to trust and risk. First of all we need to establish trust in another actor. Based on this trust, we may be able to accept a certain risk when interacting with this actor (the dashed line in Figure 1). However, since we are not able to completely validate our assessment, there always exists the possibility that the actor poses a (significantly) larger risk than what we can accept (dotted line in Figure 1).

Combining our conclusions, we summarise that:

- the compliance of insiders to policies for control and monitoring will peak and then decline—at exactly which point depends on organisational factors that require more research;
- as policies for control and monitoring increase, as expected the probability of insider threats falls;
- at some crucial inflection point, however, two events occur: first, compliance with "too many" policies starts to fall, while insiders continue to gain knowledge that makes them potential high-level insider threats. Thus the combination of these two factors (which need not be simultaneous) means that the risk of insider threats starts to increase again. Furthermore, since the high-level insider is more fully knowledgeable about the organisation, their potential for damage as an insider threat is high.

A note seems in place regarding Figure 1. We implicitly assume that the factors considered, knowledge and authentication, both evolve over time. One might argue that for many actors in an organisation the risk does not increase over time, or the trust/risk relationship does not become more complex. However, even though an employee "only" gets to know the system better, he also understands better how to perform actions that he wants to not to be ob-

served, or where to leave "markers" to document that he did something [22].

4. Policies and Compliance

We can think of policies in two basic forms:

- those that control or monitor behaviour to attempt to enforce the trust relationship (*e.g.*, through access control or monitoring of behaviour); and
- those that motivate insiders to "act in the appropriate way" — in other words to act in a way that ensures that they do not become insider threats.

In this section we will consider the impact and economics only of the first sort – those that seek to control behaviour. As trust relationships grow more complex we observe distinct differences in the economics and effectiveness of these sorts of policies.

To account for the difference between trust, acceptable risk, and potential risk as described in the previous section, we use policies to control the admissible actions, and the accessible assets. The goal of these mitigating factors clearly is to minimise the likelihood of a big differential between acceptable and actual risk or threat.

All restrictive policies seek to control or monitor behaviour. The costs of these policies, especially the hidden costs of policies interfering with the normal work flow of the organisation, can be high. This cost, however real, may be difficult to measure. Gaps and conflicts in policies can create confusion among insiders in terms of "what is right" or "how do I get my job done?" While ideally security should support people in doing their jobs, several examples are known of technological security approaches that, because they interfered with the 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; motion detectors designed to automatically log off users were disabled by covering them with plastic cups). Compliance with security policies is hard. Making compliance easy for insiders is absolutely necessary for any successful effort to constrain insider threats. Yet none of these instances lend themselves to clear-cut cost measurements, but intuitively they cost the organisation if not in money then in factors like staff time or motivation.

4.1 Enforcing Simple Trust Relationships

Control of simple trust relationships lends itself to access control and monitoring policies with commonly acceptable cost/benefit ratios. Typical questions faced when enforcing simple trust relationships are

- Who should have access to what information?
- Under what circumstances, and how defined?

We would posit that, although restrictive policies have organisational costs, some of these measures appear to have acceptable cost/benefit ratios. Basic access control measures (passwords or tokens, required and automatic cryptographic use, selective file access) and monitoring (to a point) appear beneficial in preventing or discouraging a large set of insider threat activities that could create a potentially large loss to the organisation. The deciding factor in all of these cases is how much monitoring and access control is acceptable (both ethically and legally) and at what point does it stop being beneficial, compared with the costs (both monetary and otherwise) to the organisation.

It also appears that it is commonly understood that there is a "reasonable" probability that these measures will prevent certain common types of insider threats. None of this is supported, to our knowledge, by anything other than anecdotal evidence.

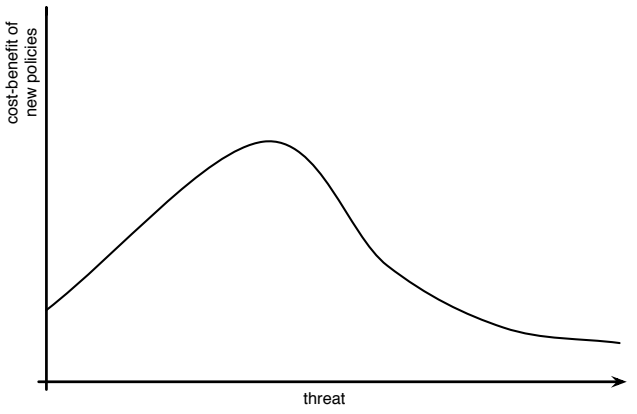

Figure 3. Plot of cost-benefit ratio of new policies against an increasing threat.

The impact of these policies is aggregative up to a point—in other words, certain sets of policies work together to create a greater benefit compared to cost than they would individually. For instance, passwords together with physical limitations on data copying (blocking certain ports, for example) together with selective monitoring together may provide much greater benefit than that provided by each policy separately. Part of the reason for this, simply, is that policies controlling simple trust relationships oftentimes affect a large number of insiders (passwords may be required for all insiders, for example), and that to a point combinations of these policies reinforce each other.

The marginal effectiveness of each additional policy declines, all other things being equal. Thus, up to a point we conclude that the probability adjusted benefit-cost value for restrictive policies aimed at insider threats is positive, and may even be increasing, up to a point. In Figure 3 we illustrate this argument. Up to a point we are able to predict that new policies will benefit the organisation, based on a reasonable risk analysis—our actions to reduce insider threats do more good than harm to the organisation. Beyond that point, however, the added policies harm the organisation, either because employees do not comply, or because they disturb the work flow too much [3, 4].

An important caveat is worth repeating—none of the factors going into this evaluation have, to our knowledge, any sound basis in data; nonetheless our description above captures (albeit in somewhat different language) a set of sentiments commonly expressed by practitioners dealing with insider threats.

4.2 Managing Complex Trust-Risk Relationship

Policies for controlling complex trust relationships face a number of challenges not faced to the same extent when controlling simple trust relationships.

The questions in terms of controlling complex trust relationships include those mentioned for simple relationships, plus:

- When does complex behaviour signal that an insider threat is taking place, as opposed to, say, creative activity?
- The effectiveness of a particular policy is unclear—are we putting in place policies that deal with potential threats that never will materialise?

Just like simple relationships, policies for controlling complex trust relationships face a number of challenges.

The aforementioned cost of policies interfering with the natural work flow of the organisation increases, we posit, for large complex systems with equivalently complex trust relationships. In

such cases prevention and detection require a significant effort. Not only may expanded monitoring (e.g., anecdotally many professionals object to the notion that their use of the computer is being monitored) affect trust within an organisation, it becomes increasingly nuanced in what to look for in complex trust relationships. For example an inordinate amount of system searching may indicate that an unauthorised person has access to that account (a masquerader)—or a forgetful mind. The cost of false positives may be significantly higher for senior managers than for data entry clerks—or even IT administrators.

Solutions may themselves be complex, and have limited applicability across the organisation. For example a set of actions to ensure that senior executives do not steal vital information in order to create their own company or move on to a competitor requires, at a minimum, heightened monitoring of system activity. But if the data is commonly used, and commonly used by the staff of senior executives, then the problem of actually detecting data theft might become immensely intrusive both to the ability of the staff and senior executives to do their work, and to morale and other human factors. A threat which may be of immense impact if it happens, but of totally unknown likelihood, and affecting only a very small number of insider s directly, probably only has high cost solutions to preventing it – if it has any at all.

Attempting to control complex trust relationships increases the risk that those actions will severely damage the organisation.

As noted above, restrictive policies (monitoring, access control) all carry the risk of increasing the cost to the organisation by interfering with people's ability to do their job. We find it intuitive that attempting to control complex trust relationships carries with it an especially high cost—in fact one that may not be acceptable to the organisation.

Thus, organisations attempting to manage the risk of high-level insider threats face a number of special challenges:

- The probability of a high-level insider threat event is difficult (impossible?) to predict or even imagine in advance.
- The longer an individual is in the organisation (or some other descriptor that captures this notion of increasing trust), the greater their knowledge of the valuable information assets or services and how to circumvent the policies in place.
- So a trusted person is also in the position to do the most damage to the organisation.
- The cost of more information security policies is poorly understood, but in general the anticipated cost is if anything less than the real cost (in other words, a well meaning set of policies runs the risk of damaging the organisation severely and in unanticipated ways, but it is unlikely that the real cost is far less than what was anticipated).

The difficulties in dealing with insider threats are increased by the complexity of organisational and insider threat goals, which we now discuss.

4.3 Simple vs. Complex

The boundary between simple and complex insider threats is blurry. By "simple" insider threats we mean those that are obvious, like the linear dependencies in [18]; while they might potentially cause severe damage, they can easily be identified and monitored. This might for example be the confidential document where every insider with access rights might pose a threat. When considering policies this would typically involve actors, roles and assets that are mentioned explicitly in policy rules.

Complex insider threats, on the other hand, got their name from Perrow's complex dependencies [18]. Here it is often unclear how they built up over time as a combination of different factors discussed above. These threats may develop "under cover", and eventually be triggered by apparently unrelated events, which exactly makes them so hard to predict.

5. Organisational and Insider Goals

Goals shape what is important to both the organisation and the insider; goals also shape what options are chosen both by the organisation and by the insider.

5.1 Organisations

Organisations have many, potentially conflicting goals that also influence how they choose to deal with insider threats. Most important they of course try to maximise their gain function, most often in the form of maximising the organisation's profit. This is supported by trying to minimise the risk of both outside and inside attacks. Factors in reaching these goals are trying to ensure (maximise) compliance with the organisation's policies as described in Section 4, to try to maximise the employee's loyalty with the organisation, and to find the right number of policies.

Poorly articulated and conflicting goals make it more difficult to determine both what is of value to the organisation, and what trust relationships in the organisation are most critical.

One key question is whether organisations who have suffered insider threats now act differently than they did in the past. And, what they are prepared to pay to avoid another occurrence? In other words, do organisations "learn" over time or by experience so as to forge clearer links between their goals and the most important values and trusts? Anecdotally, past insider threats seem to raise awareness of the threat, but it is unclear whether this also leads to more effective measures. To preview our conclusions, for insider threats that violate highly complex trust relationships the specific threat may be strictly unique.

Organisations that have faced high-level insider threats before probably do act differently. However, the only truly effective responses are not controls—how can the next high-level insider threat be anticipated? The effective responses are to first help build the organisational culture where insiders do not want to become threats, and second, to consider ways in which damage can be mitigated after the fact.

5.2 Insiders

Insiders have complex, poorly articulated goals, too—they want to, e.g., maximise the damage to the company/CEO/... or their personal gain, at the same time trying to minimise the risk of being detected. Just like organisations, insiders often have muddled goals, and organisations cannot completely predict the many forms that insider threats might take. If, as we argue, the high-level insider also has a strong incentive to be creative in their threat, then predicting in advance the form of the charismatic threat becomes even more difficult—indeed, we might conclude, almost impossible.

6. The Risk of Risk Analysis

In particular more complex trust relationships pose a set of difficult questions when performing risk analysis.

As noted above, we observe that complex trust relationships generally are associated with more complex behaviours. Thus, understanding the nature of the threat itself in any actionable way, the potential losses accruing, and the probability of such instances happening (even if they can be imagined beforehand) are all difficult.

Major, complex insider threats appear to be rare, and largely unique in their construction and execution. Of course, successfully executed, their impact on the organisation can be very large, and they should therefore be accounted for in the risk analysis. There does, however, not seem to be an adequate way of systematically

deciding that "this potential complex threat is more likely that that threats", nor any generally accepted perception across the community such as exists for less complex insider threats.

Since a priori it is difficult to predict the form that a high-level insider threat will take, organisations cannot adequately anticipate beforehand the possibly high costs that insider threats could have to their systems and enterprises. We hear frequently from corporate managers that they did not appreciate the value of what was lost through the insider threat until after the event. More formally, with poorly articulated goals making it difficult at best to estimate the value of organisational resources, and possibly highly complex insider threats affecting many different organisational resources, of course organisations have great difficulty in anticipating the costs of some insider threats.

Estimating losses from high-level threats is also challenging, since it frequently does not show up until some time after the event started. A currency speculator may appear for years to be a major profit centre for his banking group—until one day he is discovered to be an insider threat. Or, the loss is virtual until it hits. It could even be that maybe the risk is constant, but due to the risky behaviour going on for some time, the disastrous effect is getting bigger and bigger.

As described above, at the same time risk builds up in the background. Thus, complex trust relationships develop over time. In some cases this simply may be due to greater familiarity over time with the workings of the information system or in their daily work; for example over time an employee may learn or be able to guess the passwords of fellow workers. In other cases the trust relationship that extends over time is more complex. The important observation, however, is that over time more complex trust relationships grow between insiders and others in the organisation.

It is not a problem, until high-level or charismatic insiders go bad and use that knowledge to maximise their goal. We think it therefore is crucial for mitigation to make this risk explicit in an organisation's risk assessment. But just like insiders often are able to do harm because they know the system and can play it, the same holds if they are aware of what the risk function looks like.

Thus, in parallel the consequences of violating those trust relationships can become more costly to the organisation. As a gross generalisation, certainly not always true, insiders have the potential to cause more damage to an organisation the longer they have been an insider, simply as a function of the greater trust relationships that may have been established.

6.1 Plotting the Value Function

As stated above, we consider two different situations; either the organisation can anticipate a type of threat, or it can not even imagine it.

For the first case it seems that the value function of the organisation will be convex—in other words up to some point we can anticipate the most common (or imagine that we can anticipate...) types of threats; the policies put in place are not too costly; we perceive their effectiveness as being high; and we believe that the probability of these types of threats to be high enough to worry about.

Thus, up to a point the value function of the organisation looks as described before (Figure 3).

This assumes almost perfect information—we can anticipate a certain type of threat, though we do not know who will emanate the threat; we can estimate its probability of taking place; we know the cost of putting in place policies to address this threat; we know how effective these policies will be, *i.e.*, the probability that they will prevent or detect an insider threat; we know what the cost to the organisation will be.

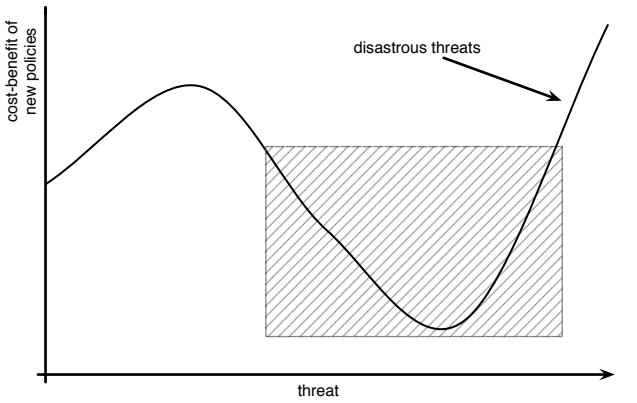

Figure 4. Extension of the cost-benefit ratio plot against threat. Once the threat reaches disastrous levels, it may very well be beneficial to instantiate policies trying to prevent these events. Organisations face the problem that in a certain area they are unable to predict the effect of policies on optimising their overall gain function—this is exactly the area where senior managers throw their hands in the air and choose to ignore the threat.

It seems logical that in this case there will be some very serious (but not totally absolutely catastrophic) insider threats for which the cost will exceed the benefit of putting in place the required policies adjusted for their likelihood of being successful. This complex nature of threats causes a decrease in the cost-benefit ratio as illustrated in Figure 4, weighting cost of policies against their probability weighted value, since the added policies have a negative effect on compliance, or work flow, or a combination thereof.

As the impact of threats increases, there can occur some absolutely disastrous threats, whose outcome is absolutely unacceptable for the organisation. For these, the probability-weighted value is positive again, as it is inevitable that these threats are mitigated—in other words, it is appropriate to take action.

But in a more "real" circumstance we observe that once the organisation gets to a certain point—beyond the "handful" of normal actions that we would take (whatever those are; access control, monitoring, periodic background checks)—what we have entered is totally unknown territory even if we can anticipate the nature of the inside threat. In this area we do not know how our employees react to more policies, how our gain function evolves, how the risk of attacks evolves, and so on. This unknown territory is marked by the box in Figure 4.

This unknown territory is defined by:

- At some point we start to get on shaky ground in terms of estimating the true cost to the organisation of the policies— *i.e.*, insiders get irritated with the working environment, or we start to increase the risk that in some sort of unexpected situation needed data is not accessible;
- Additional policies may also increase the likelihood of false positives, or also the cost of false positives goes up— *e.g.*, as we monitor senior executives; and
- As we move further out on the trust curve, we get less confident that the threat we are trying to solve is real (or is it just imaginary?). But the cost of the additional policies is real.

Thus the organisation's real value function looks like that in Figure 4, with the boxed area replaced by a huge question mark. This is exactly the area where senior managers state that they throw the hands up in the air, being aware of the threats, but also being aware of not knowing how their organisation will behave. And, while they

are at it, they often ignore the high-risk threats as well, taking them into account, because, *e.g.*, the risk may be high, but so is the gain. Besides the Binney case, all examples mentioned above fall into this category. Kerviel was earning his bank huge amounts of money before going bad, so it might have been convenient to ignore the risk, and in the case of the tax fraud, the insider's suggestion not to implement a certain auditing system was followed since the budget had already been overspent.

6.2 The Benefit of Obscurity

It should be noted that a risk analysis itself, once performed, poses a significant risk to the organisation; this is especially true if we consider higher management as potential insiders. Since they certainly have a complex risk/trust relationship to their organisation, it seems at least mandated to do so.

Once a detailed risk analysis has been performed, it may be hard to keep secret, *especially* from upper management. Ironically, the very risk analysis that is performed to identify and *limit* the effect of insider threats (or threats in general), does actually *increase* their potential effect if the result gets in the wrong hands. The same information, being confidential, is much less harmful in relation to outsiders, and the threat they pose will therefore not increase.

We argue therefore that for the result of a detailed risk analysis Kerckhoffs principles [15] should *not* be applied, since its content can cause disastrous damage and should therefor be accessible only to a very limited group of actors. This, however, may lead to a circular dependency, since the risk analysis may be needed to identify who should be allowed to access its results.

It should be noted that this approach of "security by obscurity" might also seem advisable for selected other documents, which could be described as the spinal cord of a company. However, it might be infeasible to identify who can or cannot be trusted to access these documents. Eventually one has to trust actors to behave well.

7. Strategies to Change Motivation Rather than Prevent Bad Insider Actions

This points to organisations behaving economically rationally for all but high-level threats by picking a small number of insider threats that can be managed, and dealing with the rest through mitigation after the fact. There may be some threats posing such a great risk to the organisation that the cost to the organisation of the necessary policies may be justifiable. However, most high-level threats are, by nature, unpredictable.

For high-level insider threats, two other types of policies may be most useful:

- Mitigation of the impact of the insider threat. Are there ways of increasing the successfulness of mitigation? Are there types of insider threats for which mitigation is just not going to be an acceptable path? We suspect that the potential damage from a high-level insider threat may be too great to think of mitigation as a relief (for example, the case of Aldrich Ames, the insider who spied on behalf of enemies of the United States, does not appear to lend itself to mitigation).
- However, investment in the other sorts of policies—changing behaviour so that people trust their organisation and do not *want* to cause harm—makes the most sense. Even though these sorts of "positive" policies are even less well understood in terms of their effectiveness/impact than the technically based "control" measures which we show break down at a certain point, anecdotes suggest that friendly, supportive, organisational cultures, where insiders do not have the incentive to become a threat, are possible to construct. Even difficult situations, like a large number of firings, can be done in a way that preserves a positive atmosphere.

8. Conclusion

We conclude, therefore, based on this logic, that risk analysis for insider threats is useful up to a point, but that the whole risk analysis approach as a means of selecting what actions to take breaks down as we get into the territory of dealing with highly complex trust relationships—insiders who are highly knowledgeable about the information, its value, and the protections in place. We can imagine all sorts of threats, but do not know which ones to take seriously. Maybe too as we get into highly specialised threats the types of policies we would take to counter each threat become less universal, and more specialised.

We see the net effect of risk analysis breaking down in all sorts of organisations. This article began by noting that organisations act as though they tolerate some serious insider activity—in other words, that in addressing the insider threat there is an even worse perceived risk of severely damaging the organisation.

We also observe organisations figuratively throw up their hands in the face of a threat that, while recognised, seems impossible to adequately address. Consider for example a complex organisation like a hospital. Even *defining* a trust relationships strikes us as being very difficult, time consuming, and prone to errors. Having defined (somehow) the trust relationships at risk of an insider threat, the organisation is still faced with the task of developing policies to counter the threat. Is it any wonder then that some organisations throw up their hands in the face of this challenge?

8.1 Probability of policies being successful in blocking high-level insider threats

To further our conclusion, we note that all of this is that policies have a probability of being successful (that they actually work). So the expected loss function is the *probability of an insider threat to occur*, times the *probable damage of a certain amount or type*, times the *probability that the policies imposed will be unsuccessful* in blocking that threat. We believe that for more complex trust relationship based insider threats the very effectiveness of the policies deployed to counter the threat may be less effective. This goes in line with observations that organisations with increased surveillance and auditing often state that the number of detected cases stays constant, as was recently reported by several public agency and private company officials [20].

So for high-level insider threats it is very expensive to put in place all of the policies to block these threats, with increasingly low probability that the policies will actually be successful (because the more policies you add the less successful cumulatively they will become). The loss function is very high at one end, with low probability throughout, but when they do occur it's a big loss.

To summarise: our chief tool for assessing threats (risk analysis) and for deciding what threats to deal with, and how, breaks down for what might be the worst sorts of threats. This finally explains why organisations behave as they do, and that, even though surprising, their behaviour is economically rational even in the face of high-level threats—by picking a small number of insider threats that can be managed, and dealing with the rest through mitigation after the fact (even if mitigation is not likely to be very successful). For high-level threats it may be that in a few cases (where the event can be anticipated in advance, and the costs to the organisation are very high) the organisational cost and disruption of imposing control policies may be worth it. But this probably describes the exception rather than the rule.

The appropriate insider threat control strategy depends on an organisation's perceived loss function from insider threats. Different organisations presumably have differently shaped loss functions:

US intelligence organisations probably have a big bump at the far right, making them very sensitive to high-level insider threat. Banks are probably like intelligence organisation, though the evidence is mixed on this.

We conclude as well that it becomes economically rational at some point in the threat function to invest heavily in policies to change behaviour in a positive fashion, even if these policies are not well understood in terms of their impact or effectiveness.

References

[1] A. Adams and M. A. Sasse. Users are not the enemy. *Commun. ACM*, 42(12):40–46, 1999.

[2] R. H. Anderson. *Research and Development Initiatives Focused on Preventing, Detecting, and Responding to Insider Misuse of Critical Defense Information Systems: Results of a Three-Day Workshop*. RAND Corporation, Santa Monica, CA, U.S.A., 1999.

[3] A. Beautement, R. Coles, J. Griffin, B. Monahan, D. Pym, M. Sasse, and M. Wonham. Modelling the human and technological costs and benefits of usb memory stick security. In *Proceedings of the Workshop on Economics in Information Security*, 2008.

[4] A. Beautement, M. Sasse, and M. Wonham. The compliance budget: Managing security behaviour in organisations. In *New Security Paradigms Workshop*, 2008.

[5] Binney v. Banner Therapy Products, 631 S.E. 2d 848, 850. North Carolina Court of Appeals, 2006.

[6] M. Bishop, S. Engle, S. Peisert, T. Whalen, and C. Gates. Case studies of an insider framework. In *Proceedings of the 42nd Hawaii International Conference on System Sciences (HICSS)*, 2009.

[7] P. Cofta. *Trust, Complexity and Control: Confidence in a Convergent World*. John Wiley and Sons, 2007.

[8] E. Cole and S. Ring. *Insider Threat: Protecting the Enterprise from Sabotage, Spying, and Theft*. Elsevier, 2006.

[9] *Computer Crime and Security Survey*. Computer Security Institute, 2007.

[10] B. T. Contos. *Enemy at the Water Cooler*. Elsevier, 2007.

[11] J. Hunker and C. Bulford. *Federal Prosecution of Insider Threats Demonstrates Need for Reform; Analysis based on data base of Federal prosecutions since 1995*. Manuscript under review, February 2009.

[12] J. Hunker, J. Predd, S. L. Pfleeger, and C. Bulford. Insiders behaving badly: A taxonomy of bad actors and their actions. *Manuscript under review*, 2008.

[13] Jérôme Kerviel. Available from `http://en.wikipedia.org/wiki/Jerome_Kerviel`, last visited February 27, 2009.

[14] D. Keating. Tax suspects guidance on software left d.c. at risk. Washington Post, June 2008.

[15] A. Kerckhoffs. La cryptographie militaire. *Journal des sciences militaires*, IX, 1883.

[16] F. H. Knight. *Risk, Uncertainty, and Profit*. Hart, Schaffner & Marx; Houghton Mifflin Co., 1921. Library of Economics and Liberty [Online] available from http://www.econlib.org/library/Knight/knRUP.html; accessed 24 May 2009.

[17] M. Michelson. Bank scandal a blow to french pride. In *International Herald Tribune*, January 2008.

[18] C. Perrow. *Normal Accidents: Living with High-risk Technologies*. Princeton University Press, 1999.

[19] J. Predd, S. L. Pfleeger, J. Hunker, and C. Bulford. Insiders behaving badly. *IEEE Security and Privacy*, 6(4):66–70, 2008.

[20] C. W. Probst, J. Hunker, M. Bishop, and D. Gollmann. Countering insider threats. *Dagstuhl Seminar Proceedings*, 2008.

[21] G. Schudel and B. Wood. Modeling behavior of the cyber-terrorist.

[22] E. E. Schultz. A framework for understanding and predicting insider attacks. In *Proceedings of CompSec*, 2002.

[23] Stein Bagger. Available from `http://en.wikipedia.org/wiki/Stein_Bagger`, last visited February 27, 2009.

[24] D. Weirich and M. A. Sasse. Pretty good persuasion: a first step towards effective password security in the real world. In *NSPW '01: Proceedings of the 2001 workshop on New security paradigms*, pages 137–143, New York, NY, USA, 2001. ACM.