From Insider Threats to Business Processes that are Secure-by-Design*

Dieter Gollmann†
Hamburg University of Technology
Hamburg, Germany
diego@tu-harburg.de

Abstract

We argue that insider threat is a placeholder term that accompanies the transition from securing IT infrastructures to securing the socio-technical systems made possible by these IT infrastructures. The term insider in its literal interpretation loses meaning in a context where there are no stable perimeters one can refer to. Business practices such as outsourcing, employing temporary contractors, and the very use of IT, have removed security perimeters in the search for short-term efficiency gains, which may result in mid-term losses due to increased vulnerabilities. We conclude that securing socio-technical systems calls for the design of organisational (business) processes that remain viable once inside information about their implementation becomes available to potential attackers rather than for the deployment of secure IT infrastructures.

Keywords: Insider threats, business process, IT security

1 Introduction

Security is a fashion industry. Ever so often a new buzzword gets adopted that dresses up issues that had been around for a long time. Insider threats are an example in case. This term has entered discourses about security in recent years, reaching also into the research community, see *e. g.* [1, 2, 3]. However, the observation that threats to an organization can emanate from insiders is old news in IT security. Back in the early 1990s depending on whom you asked, 70%, 80%, or 90% of security incidents in commercial entities were due to insiders, where counts may have referred to the number of incidents or to the size of the damage.

We cite an incident from 1966 documented in [4] to support our case that insider threats are not new at all. A programmer at a bank had set an unlimited overdraft on his own account. This was discovered when the computer broke down and accounts were processed manually. The perpetrator was dismissed, taken to court, received a suspended sentence, paid back the overdraft, and was re-hired as contractor. This attacker was an insider in two ways, as a developer of the system and as a customer of the bank.

Insider threats may have become less prominent an issue—briefly—when companies and customers went on-line from the mid 1990s onwards. Public perception of IT threats, moreover, was changed when hacker stories became the myth of that age, amplified by blockbuster movies like Independence Day (1996) and front page news about worm attacks in the mass media.

The realization that insiders can pose threats to an organization is not new in business and predates the age of IT by millennia. The most pernicious adversary, and the felon punished most severely, is the traitor within. When we look at the ways organizations have protected themselves we may heed the familiar warning "security is not an add-on feature". This comment has been made in the context of

Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications, volume: 3, number: 1/2, pp. 4-12

*Research on this paper was in part supported by the BMBF project ContainIT under grant number 13N11013.

†Corresponding author: Harburger Schloßstr. 20 (HS20), 1st floor, room 123, 21079 Hamburg, Germany, Tel: +49 (0)40/42878-3026

many an IT system, *e. g*. on the design of operating systems, on mobile phone networks, and more. It equally applies to other systems and processes that manage the assets of an organization.

We will start our investigations into the nature of the IT security challenges faced today with a brief overview of how concepts from the business domain have entered IT security. We then pursue the argument that drawing meaningful boundaries so that a business process is executed entirely within such a boundary is difficult in today's dynamic heterogeneous organizations. Hence, the concept of an insider is not helpful when designing security solutions. We then focus on the processes themselves. This in turn will lead us to a preliminary discussion on risk analysis for socio-technical systems.

1.1 IT Security and Commerce

Business has centuries of experience in dealing with dishonest staff and business partners. This experience was first brought into IT security with the Clark-Wilson model in 1987 [5], one of the first major contributions to IT security not related to research driven by US defense. The authors came from the MIT Laboratory for Computer Science, Clark, and from the accountancy firm Ernst & Whinney, Wilson. The Clark-Wilson model takes it cues from commercial practice and puts a strong emphasis on integrity.

- Internal consistency refers to properties of the internal state of an IT system.
- External consistency refers to the relationship of the internal state of an IT system to the outside world.

The main mechanisms for maintaining integrity are well-formed transactions and separation of duties. In the security incident narrated above we can see a violation of a separation of duties rule. Should a developer of a banking software be a customer of that bank at the same time?

The separation between a front office making deals and a back office clearing the deals in investment banking is an instance of separation of duties intended to limit the risk a single trader can expose a bank to. Nick Leeson managed to circumvent this control when bringing down Barings Bank[1][2] in 1995. This case is analysed from an IT perspective in [6] and from the perspective of sociology in [7]. It again involves software installed by the insider that hid the traces of unauthorized transactions and an event outside the insider's control that caused the fraud to be exposed, in this case the Kobe earthquake on January 17, 1995. Similar observations apply to the case of Jérôme Kerviel (2008) [8].

The Clark-Wilson model places *transformation procedures* at its core. They serve as an intermediate layer between the subjects and objects of access control. For each data type, applicable access operations (transformation procedures) are defined, while users are authorized to perform certain transformation procedures. A subject (process) gets access to a data item based on the user it is speaking for and the transformation procedure being executed. The presentation of the model puts a strong emphasis on the proper certification of the transformation procedures an organisation decides to implement, *i. e*. on the management decisions confirming that automated IT procedures are in line with organisational processes.

There is a close conceptual link from the Clark-Wilson model to role-based access control (RBAC) [9, 10]. A role is defined as a collection of (transformation) procedures. Users are assigned roles and can access data only via the roles assigned to them. Academic research on RBAC has in the main explored the formalization of access control models, *e. g*. models with role hierarchies or models for administrative RBAC. Important challenges in the real-life application of RBAC relate to role-engineering [11, 12], *i. e*. to defining meaningful functional roles in dynamic business environments. Important challenges furthermore relate to the design of the business processes themselves. This is the topic we will now turn to.

[1] http://www.numa.com/ref/barings/bar03.htm

[2] http://www.riskglossary.com/link/barings_debacle.htm

2 Business Processes

Transformation procedures, or well-formed transactions, automate (parts of) organisational processes. For brevity, we will be referring to business processes but our considerations do not only apply to commercial entities. Executing a business process incurs costs. It is a natural objective for an organization to minimize these costs. Indeed, this is a major rationale for automating processes using IT.

Business processes may be subverted by parties involved in their execution. These parties may justifiably be considered as *insiders* with respect to the business process. A business process may also be subverted by external parties that attack the IT infrastructure used to execute the process. This can be done either by manipulating documents exchanged during the process, or by compromising one of the hosts involved.

Losses due to attacks add to the total costs of executing a business process. Implementing security controls in a business process incurs costs. We would thus be facing a standard optimization problem if we could properly quantify the cost of defences, the impact of defences, the potential damage caused by attacks, and the likelihood of attacks. In general, we are lacking strong empirical evidence and have to work on the basis of assumptions that are more or less plausible. We will return to this topic in Section 3.2.

2.1 Business Processes in the Web

When a business process is novel, insider knowledge is limited and attacks are rare. Security controls may still be lacking as there is no known threat to defend against. Initially then the new business process appears to be more efficient. When the attacks start the picture changes. This should not come as a surprise! (In)security costs not visible in the short term can manifest themselves later on. At that stage, the clamour for better security may arise and processes may be "strengthened" by adding some security features.

In such situations, security would be appear to be an add-on feature after all. Setting up secure tunnels for communications or hardening operating systems can be done independently of the business processes, at least to some extent. Cryptographic protocols, for example, require their own infrastructures for key management, which have to fit the business processes to be secured.

The Web has been increasingly used for re-engineering business processes in the last decade. In such efforts, short-term efficiency gains can be made by neglecting security controls[3]. For a time, the main threat in the Web came from so-called "hackers" launching sweeping attacks against the entire IT infrastructure but not targeted attacks against specific business processes. The situation reversed from around 2004 when traditional criminal activities, viz attacks for financial gain, became the dominant issue.

2.2 Insider Threats

The attention paid to insider threats today can be explained from the observation that mechanisms for protecting the IT infrastructure were by and large designed with external attackers in mind. These mechanisms may thus be ill suited for attacks coming from authorized parties or from other persons with inside knowledge of a business process and its implementation. Incidentally, the dealings leading up to the collapse of Barings Bank draw a fascinating picture of how various insiders were unwittingly or negligently colluding in the attack [7].

[3]There is an analogy to software engineering. Some optimizing compilers remove apparently redundant checks from code for better performance.

It would then seem a plausible strategy to complement traditional IT security controls by identifying the various insiders, their motives, their attack options, and the countermeasures available. This argument has been examined comprehensively in [3], reaching the conclusion that "the insider threat is not one problem, but many". We follow the line taken in [13] on *trust*: a term with too many meanings will confuse rather than lead to the solution of the underlying problem. Overloaded terms are *placeholders* indicative of the fact that a field has reached the point where established concepts are insufficient but where the new concepts have not been formed yet that would allow to address the novel challenges encountered.

What is then the new challenge corresponding to the placeholder "insider threat"? Arguably, we are seeing the transition from securing the IT infrastructure to securing socio-technical systems. IT security is concerned with effects in the IT domain; in communications security, for example, the job is done once a message is delivered securely; concerns about what may go wrong when the message is processed further are out of scope. Socio-technical systems have effects in the social domain. As a general defence strategy, processes are constructed in a way so that a successful attack has to compromise multiple points in the system. Separation of duties is one example for this strategy. *Consistency checks* between observations made at different points in the system can detect interference with a process. Such checks can be performed automatically in the IT system, but also manually by people tasked with auditing functions.

Insider threat is the wrong paradigm for dealing with the challenge we have outlined. The term *insider* demands by necessity a well-defined security perimeter. It must be possible to talk about entities inside or outside this perimeter. This demand has to be put up against business reality: we often lack clear organisational perimeters.

- In an organization, people come and leave; temporary contractors may be employed. Thus, insider knowledge will exist outside of the organization.
- Tasks may be outsourced so processes will be executed in parts outside the organization.
- IT systems, in particular software components, are sourced externally; furthermore, IT operations may be outsourced too. Knowledge about the infrastructure, or the infrastructure itself, are outside of the organization.

Subcontracting and outsourcing supposedly lead to efficiency gains but transcend boundaries where defences could have been placed. We once more arrive at the trade-off between short term gains and long term impacts on efficiency. Note that this trade-off is often made in environments where the impacts of business decisions are evaluated within relatively short time frames.

2.3 Sustainable Business Processes

When the term *insider* and the insiders themselves are too elusive for applying security controls we may return to our original goal: how to design business processes so that the damage attacks may cause is limited? In this light, we adopt *business processes that are secure-by-design* as our paradigm. We will call those processes sustainable to capture the fact that they remain viable once insider knowledge about their implementation is available to attackers.

> **Definition**. A business process is *sustainable* if its total execution costs remain at an acceptable level even when attacks are launched with insider knowledge.

When business processes change, attack goals may change. For illustration we point to an incident[4] that occurred during the privatisation of British Rail in 1996. As background, travelcards in London were valid on British Rail and on London Transport lines (buses, underground). Revenue from travelcards was shared according to estimated journey patterns; for tickets sold at Fenchurch Street (no underground) London Transport received a share of 22% of ticket sales but for tickets sold at Upminster (underground station) a share of 48%. Management of the Fenchurch Street to Tilbury line had travelcards printed at Fenchurch Street and sold from their ticket offices in Upminster to increase the profitability of their operations.

There was no incentive for such an action before the advent of privatisation. This practice was detected when internal audit at British Rail noted a steep increase in ticket sales at Fenchurch Street. The management buyout of the Fenchurch Street to Tilbury line was stopped at the last moment.

3 IT Security Basics

Organizations need to secure their assets. Assets are manipulated by business processes. When designing business processes we must therefore perform a risk analysis for these processes and for the socio-technical systems executing these processes. Such a risk analysis can follow well established patterns, see *e. g*. the ISO 27000 series of standards[5].

- Model the system and identify the assets.
- Identify generic vulnerabilities.
- Identify generic countermeasures.
- Identify specific threats based on the assets, on the actions the system is designed to perform, and on the actions assumed to be available to attackers.
- Rate likelihood of threats,
  - by consulting experts,
  - based on statistical evidence.
- Compute risk from the likelihood of threats and value of assets, rank countermeasures.

The challenges lie in the concrete instantiation of this general strategy. We will comment specifically on three aspects of this instantiation, modelling socio-technical systems, uncertainty, and return from divide-and-conquer.

3.1 Modelling Socio-technical Systems

Models for socio-technical systems should capture the *interfaces* between the IT systems and the physical world. At this point we can also capture data flows that support consistency checks, which are an important feature of well-formed transactions. When compiling a model it is advisable to start at a more abstract level but certain attacks will not be visible at higher abstraction layers. The modelling methodology should thus facilitate views on the system at several layers of abstraction, *i. e*. *vertical* decomposability.

[4] http://www.independent.co.uk/news/rail-fraud-aimed-to-help-success-of-selloff-1317413.html
[5] http://www.27000.org/

With complex business processes obtaining a complete picture may be onerous. Parties may only have local views of the processes they are involved in, a situation we encountered when modelling logistics processes for the container transport. The overall model has thus to be assembled from individual snapshots. The modelling methodology should thus also support *horizontal* composability of partial models.

3.2 Uncertainty

The former US Secretary of Defense, Donald Rumsfeld, once famously referred to the known unknowns and the unknown unknowns[6]. In risk analysis, the unknowns are the attackers' actions. Some attack patterns may be known but there are also novel attacks and zero-day exploits. Not preparing for the known unknowns is inexcusable; business continuity planning and incident response planning are part of responsible security management. However, we can never be fully aware of the unknown unknowns, *i. e*. all the routes an attacker may take.

There is thus a degree of uncertainty about the nature of the attacks that may occur. There is also uncertainty about how frequently a potential attack will occur, if at all. We may try to capture this uncertainty by *probabilities*. Probability has precise definitions in mathematics; there is Laplacian (frequentist) and Bayesian (subjective) probability. There have been attempts to use the latter approach for measuring the security of a system by observing the mean time between (new) attacks, see *e. g*. [14, 15]. However, in many cases risk analysis has to deal with inputs where this mathematical framework cannot be applied. In such situations *likelihood* may be used as a more informal term.

There is a "scientific" approach for calculating probabilities or likelihoods that derives these values from statistical evidence. For example, one could assume that some family of distribution functions is a plausible basis for modelling the occurrence of the events of interest. The actual parameters of the distribution function are then estimated from recorded evidence. This method links in well with other areas of risk management but needs appropriate data collections. This is old news once more. Comprehensive data collections for security incidents have been on the wish list of security researchers and security professionals for a long time. Such data collections would put us in a position to handle well-known unknowns, but is this good enough for security? To quote [16]:

> Security risk is not measurable, because the frequencies and impacts of future incidents are mutually dependent variables with unknown mutual dependency under control of unknown and often irrational enemies with unknown skills, knowledge, resources, authority, motives, and objectives ...

There is a second method for estimating the likelihood of threats: interviewing experts with the help of questionnaires. This may seem a dangerously arbitrary process, in particular when divergent advice is received from different experts. We thus desire that answers should be *repeatable*, *i. e*. the same person should give the same answer if asked again later, and *reproducible*, *i. e*. other experts should also give the same answer. Questions must be phrased carefully so that we can come close to these goals. It is usual to pursue a divide-and-conquer strategy, asking many simple questions where answers are selected from a pre-defined menu. Here, the quality of advice on how to pick the answers is of crucial importance.

3.3 Return from Divide-and-Conquer

Divide-and-conquer splits the overall question "how secure is this process" into many sub-questions. From the answers to the sub-questions, a single final risk rating is computed. We refer to this step

[6] http://www.rumsfeld.com/about/page/authors-note

$y=4$	2	3	4	4
$y=3$	2	2	3	3
$y=2$	1	2	2	3
$y=1$	1	1	1	2
	$x=1$	$x=2$	$x=3$	$x=4$

Table 1: Example for a Mehari evaluation table.

AccessComplexity	
Low	0.71
Medium	0.61
High	0.35

Severity Rating	
0.0 - 3.9	Low severity
4.0 - 6.9	Medium severity
7.0 -10.0	High severity

Table 2: Mappings between descriptive categories and numerical values in CVSS.

as *return from divide-and-conquer*. In may cases this function is encoded as a table reflecting expert opinions. Expert opinions need to be justified, a process that may become unmanageable once a table gets too large.

The Mehari methodology [17] has developed an interesting variation to this approach, applying the divide-and-conquer principle to the calculation of the final risk value. Answers to all sub-questions are mapped to the integers from 1 to 4. Instead of a single multi-dimensional table indexed by all sub-questions, Mehari uses two-dimensional 4×4-*evaluation tables* that map two variables to one intermediate variable. Table 1 gives an evaluation table combining two variables x and y. Evaluation tables are applied iteratively until the final result is obtained. Justification of table entries now relates to small data structures and may thus become more manageable.

When computing the final risk value, one also has to keep track of the dependencies between the sub-questions asked and has to consider the relationship between answers when calculating the final risk value [18].

4 Quantitative or Qualitative Risk Analysis

Risk analysis methods are customarily classified according to the nature of the values used when computing risk. Quantitative risk analysis works with numbers, operated on by some mathematical calculus, *e. g*. probability theory. Qualitative risk analysis takes its values from descriptive categories such as high – medium – low; the inputs are typically further processed in tables.

This distinction is widely used but is it meaningful? Consider the Common Vulnerability Scoring Scheme (CVSS) [19]. The CVSS calculator[7] from NIST puts twelve questions to the evaluator. For each question, there is a menu of possible answers. For example, the choices for the field *AccessComplexity* are *Low*, *Medium*, and *High*. These descriptive categories are then mapped to rational numbers. Table 2 (left) gives the mapping for *AccessComplexity*.

The CVSS base score is calculated as an arithmetic function of the converted inputs giving an output between 0 and 10. This numeric score is finally converted back to descriptive categories as shown in Table 2 (right). In [19] this method is described as quantitative risk analysis; internally numbers get manipulated. However, the values presented at the interfaces of CVSS are taken from descriptive categories indicative of qualitative risk analysis. The mapping from inputs to severity ratings could equally have

[7] http://nvd.nist.gov/cvss.cfm?calculator

been defined directly in a table. A distinction based on the internal values used for computing risk is hence not too meaningful. The distinction should be based on the inputs to the risk calculation. The distinction would then be between

- *evidence-based* (empirical) risk analysis, based on statistical evidence, and
- *judgement-based* risk analysis based on expert opinions.

5 Conclusion

The comment that security is moving to the application layer reflects on changes in the attack patterns observed in the Web since the mid 2000s. Buffer overrun attacks on operating system functions or low level network attacks have become less prevalent than attacks exploiting browser features (cross-site scripting) or weaknesses in server scripts (SQL injection). The observations in this paper point in the same direction.

IT has permeated society to an extent that may not be visible to the casual observer, bringing about wide-reaching changes to the way organizations work. Treating IT just as an infrastructure divorced from the business domain does not do justice to the extent of these changes. In this context, we have suggested that insider threat is a placeholder term reflecting the transition from IT security to securing socio-technical systems. It is an unfortunate term as it conjures up a distinction that can rarely be made in practice. In today's webified world, good security perimeters are hard to come by. Without clear security perimeters, *insider* is not a helpful paradigm when discussing security.

When there is not much gained from looking at the *actors*, we may look at the *actions* (processes) instead. Processes cannot be secured simply by equipping their implementations with standard IT security mechanisms as an afterthought. Processes have effects in the social domain and those effects need to be understood. We must therefore move from looking at IT as an infrastructure to looking at IT as a control system. Consistency checks play an important role in controlling processes. Processes are secure by design when they can resist attacks even when details of their implementation are known to the attacker.

The case for the security of a process has to build on the results of some kind of risk analysis. A security risk analysis must deal with intentional attacks. This inherently limits the applicability of empirical methods based on statistical evidence. The risk analysis methodology employed also must capture dependencies between the actions taken by an attacker and we have to accept the possibility of novel, unknown attacks.

We conclude by posing two research challenges, modelling socio-technical systems and exploring the foundations of judgement-based risk analysis methods. Progress will be made by learning from concrete case studies.

5.1 Acknowledgments

This paper builds on discussions with colleagues at the Dagstuhl Seminars 08302 *Countering Insider Threats*, 20.07.2008 – 25.07.2008, and 10341 *Insider Threats: Strategies for Prevention, Mitigation, and Response*, 22.08.2010 – 26.08.2010. The views of the author have also been influenced by observations made within the BMBF project ContainIT.

References

[1] S. Lawrence-Pfleeger and S. J. Stolfo, "Addressing the insider threat," *IEEE Security and Privacy Magazine*, vol. 7, no. 6, pp. 10–13, November/December 2009.

[2] C. W. Probst, J. Hunker, D. Gollmann, and M. Bishop, *Insider Threats in Cyber Security*. Springer, 2010, advances in Information Security, Vol. 49.

[3] J. Hunker and C. W. Probst, "Insiders and insider threats – an overview of definitions and mitigation techniques," *Journal of Wireless Mobile Networks, Ubiquitous Computing and Dependable Applications*, vol. 2, no. 1, pp. 4–27, March 2011.

[4] A. Norman, *Computer Insecurity*. Chapman & Hall, 1983.

[5] D. R. Clark and D. R. Wilson, "A comparison of commercial and military computer security policies," in *Proc. of the 1987 IEEE Symposium on Security and Privacy (S&P'87), Oakland, California, USA*. IEEE, April 1987, pp. 184–194.

[6] H. Drummond, "Did Nick Leeson have an accomplice? The role of information technology in the collapse of Barings Bank," *Journal of Information Technology*, vol. 18, no. 2, pp. 93–101, June 2003.

[7] I. Greener, "Nick Leeson and the collapse of Barings Bank: Socio-technical networks and the 'rogue trader'," *Organization*, vol. 13, no. 3, pp. 421–441, 2006.

[8] J. Epstein, "Security lessons learned from Société Générale," *IEEE Security & Privacy Magazine*, vol. 6, no. 3, pp. 80–82, 2008.

[9] R. S. Sandhu, E. J. Coyne, H. L. Feinstein, and C. E. Youman, "Role-based access control models," *IEEE Computer*, vol. 29, no. 2, pp. 38–47, February 1996.

[10] R. S. Sandhu, D. Ferraiolo, and R. Kuhn, "The NIST model for role based access control: Toward a unified standard," in *Proc. of the 5th ACM Workshop on Role Based Access Control (RBAC'00), Berlin, Germany*. ACM, July 2000, pp. 47–63.

[11] E. J. Coyne, "Role engineering," in *Proc. of the first ACM Workshop on Role-based Access Control (RBAC'95), Gaithersburg, Maryland, USA*. ACM, November-December 1995.

[12] G. Neumann and M. Strembeck, "A scenario-driven role engineering process for functional RBAC roles," in *Proc. of the 7th ACM International on Access Control Models and Technologies (SACMAT'02), Monterey, California, USA*. ACM, June 2002, pp. 33–42.

[13] D. Gollmann, "Why trust is bad for security," *Electronic Notes in Theoretical Computer Science*, vol. 157, no. 3, pp. 3–9, 2006.

[14] B. Littlewood, S. Brocklehurst, N. Fenton, P. Mellor, S. Page, D. Wright, J. Dobson, J. McDermid, and D. Gollmann, "Towards operational measures of computer security," *Journal of Computer Security*, vol. 2, pp. 211–229, 1993.

[15] S. Brocklehurst, B. Littlewood, T. Olovsson, and E. Jonsson, "On measurement of operational security," *IEEE Aerospace and Electronic System Magazine*, vol. 9, no. 10, pp. 7–16, October 1994.

[16] D. B. Parker, "Risks of risk-based security," *Communications of the ACM*, vol. 50, no. 3, March 2007.

[17] *Mehari Risk Analysis Guide*, Club de la Sécurité de l'Information Français, Paris, France, April 2007.

[18] L. A. Cox, "What's wrong with hazard-ranking system? An expository note," *Risk Analysis*, vol. 29, no. 7, pp. 940–948, 2009.

[19] K. Scarfone and P. Mell, "An analysis of CVSS version 2 vulnerability scoring," in *Empirical Software Engineering and Measurement*, 2009, pp. 516–525.

Prof. Dieter Gollmann received his Dipl.-Ing. in Engineering Mathematics (1979) and Dr.tech. (1984) from the University of Linz, Austria, where he was a research assistant in the Department for System Science. He was a Lecturer in Computer Science at Royal Holloway, University of London, and later a scientific assistant at the University of Karlsruhe, Germany, where he was awarded the 'venia legendi' for Computer Science in 1991. He rejoined Royal Holloway in 1990, where he was the first Course Director of the MSc in Information Security. He joined Microsoft Research in Cambridge in 1998. In 2003, he took the chair for Security in Distributed Applications at Hamburg University of Technology, Germany. Dieter Gollmann is the acting editor-in-chief of the International Journal of Information Security and an associate editor of the IEEE Security & Privacy Magazine. His textbook on 'Computer Security' has appeared in its third edition.