Using Violation and Vulnerability Analysis to Understand the Root-Causes of Complex Security Incidents

C.W. Johnson
Dept. of Computing Science,
University of Glasgow,
Glasgow, Scotland.
http://www.dcs.gla.ac.uk/~johnson
johnson@dcs.gla.ac.uk

ABSTRACT
There is an increasing need for incident response to look beyond the immediate causes of security violations. For example, the US Department for Homeland Security has commissioned a number of recent reports into the ‘root causes’ of adverse events ranging from denial of critical infrastructure to barriers for security information transfer between Federal agencies. The US Department of Energy has also established the Information Security Resource Center to coordinate the ‘root cause analysis’ of security incidents. A recent report by the Harvard Business School (Austin and Darby 2003) highlighted several commercial initiatives to understand not simply what went wrong in any single previous incident but also to identify any further underlying vulnerability. A common theme in all of these initiatives is to go beyond the specific events of a particular security incident and to identify the underlying ‘systemic’ technical, managerial and organizational precursors. Unfortunately, there are relatively few established tools and techniques to support the ‘root cause’ analysis of such incidents. This paper, therefore, provides an introduction to V² (Violation and Vulnerability) diagrams. The key components of this technique are deliberately very simple; the intention is to minimize the time taken to learn how to exploit this approach. A complex case study is presented. The intention is to provide a sustained analysis of Rusnak’s fraudulent transactions involving the Allfirst bank. This case study is appropriate because it included failures in the underlying audit and control mechanisms. It also stemmed from individual violations, including the generation of bogus options. There were also tertiary failures in terms of the investigatory processes that might have uncovered the fraud long before Allfirst and AIB personnel eventually detected it.

Keywords: Root-cause analysis; Security violations; Accident analysis.

INTRODUCTION
It seems unlikely that we will ever be able to eliminate security related incidents across a broad range of public and private organizations. The continual pressures for additional functionality through technological innovation create vulnerabilities that can be difficult to anticipate or guard against. The US military describe how during Operation Desert Storm and Desert Shield, ‘perpetrators who were thousands of miles away illegally accessed dozens of U.S. military systems… sophisticated break-in techniques were employed to obtain data about U.S. troop movements, ordnance systems, and logistics…new security vulnerabilities that expose systems and networks to unauthorized access and/or deny service are constantly being discovered’ (Dahlgren, 2002). Given that it is impossible to achieve total security, it is important that organizations plan their response to those attacks that do occur. For instance, the CISCO (2003) ‘Best Practices White Paper’ on network security urges companies to collect and maintain data during security incidents. This information can be used to determine the extent to which systems have been compromised by a security attack. It can also be critical to any subsequent legal actions; “if you're interested in taking
legal action, have your legal department review the procedures for gathering evidence and involvement of the authorities. Such a review increases the effectiveness of the evidence in legal proceedings”. These recommendations reflect the current ‘state of the art’ in incident investigations. The focus is on the groups and individuals who perpetrate an attack rather than the underlying technical, managerial and organizational factors that create ‘systematic’ vulnerabilities in complex systems.

There is a growing realization that security investigations must examine the root causes of security incidents. A number of organizations already recognize the importance of this ‘lessons learned’ approach to security incidents. For example, the Los Alamos National Laboratory adopted this approach in the aftermath of a series of security related incidents involving information about nuclear weapons research. The mishandling’ of two computer hard drives containing classified information led the director of the laboratory to report to the Senate Armed Services Committee. This report focused on the individual human failures that were identified as root causes. However, it also consider the contributing factors that included the ‘government-wide de-emphasis on formal accounting of classified material that began in the early 1990s, which weakened security practices and created an atmosphere that led to less rigor and formality in handling classified material’(Roark, 2000). These and similar findings have led the US government to focus more directly on the different factors that contribute to the underlying causes of security vulnerabilities. The Government Security Reform Act (2001) transferred the Federal Computer Incident Response Capability (FedCIRC) from the National Institute for Standards and Technology (NIST) to the General Services Administration (GSA). As part of this move, the GSA was charged to identify patterns in the causes of security incidents (Lew, 2001).

Similar trends can be observed in commercial organizations, especially business consultancies. For instance, Price Waterhouse Cooper (Skalak, 2003) recently issued a brief on understanding the root causes of financial fraud. They argued that ‘the key for companies is to use a global risk paradigm that considers the root causes of financial fraud, corporate improprieties and potential regulatory malfeasance arising from different markets, and therefore different risk environments, in which global enterprises operate’. Although their focus is on the wider aspects of fraud and not simply of security, the Investigations and Forensic Services group within PWC have argued that a wider form of ‘root cause’ analysis represents a new paradigm for the investigation of security incidents. The intention is to probe beyond the specific violations of external agencies and junior staff members to look at the wider organizational problems that created the context and opportunities for these threats to be realized. Several accountancy firms in the US and Europe have adopted a similar perspective as they begin to examine the consequences of recent corporate scandals. In particular, they have looked beyond the individual (mal-)practices in particular cases. It has been argued that ‘controls, no matter how sound, can never prevent or completely limit persons in high places from circumventing controls or prevent or detect all fraud …auditors do not guarantee discovery of all fraud but provide only reasonable assurance of the absence of material fraud…there have been too many instances of fraud, transactions in excess of authorized limits, and other negative events while controls were thought to be in place or auditors present to permit acceptance of these contentions. Many factors have created the current quandary. They require clear understanding and careful response for auditors and organizations they serve to rebuild the level of public confidence previously enjoyed’ (Rabinowitz, 1996).

PRIMARY, SECONDARY AND TERTIARY FACTORS IN SECURITY INCIDENTS

The previous quotations argue that specific violations that lead to security incidents often form part of a more complex landscape of external threats, managerial and regulatory failure, of poor technical design and of operational inadequacies. Mackie (1993) uses the term ‘causal complex’ to describe this causal landscape. Although he was looking purely at the philosophy of causation, it is possible to apply his ideas to clarify some of the issues that complicate the investigation of security incidents. Each individual factor in a causal complex may be necessary for an incident to occur but an attack may only be successful if they
happen in combination. Several different causal complexes can lead to the same outcomes even though only one may actually have caused a particular incident. For instance, unauthorized trading might not be detected because of insufficient oversight, collusion or through oversight that was ineffective. It is for this reason that most security investigations consider alternate scenarios in order to learn as much as possible about the potential for future failures. In our example, an investigation might look at the potential impact of collusion even if a particular incident stemmed from inefficient oversight. These high-level arguments are grounded in Microsoft (2003) technical advice for security audits: “During security risk identification, it is not uncommon for the same condition to have multiple consequences associated with it. However, the reverse also may be true there may be several conditions that all produce the same consequence. Sometimes the consequence of a security risk identified in one area of the organization may become a risk condition in another. These situations should be recorded so that appropriate decisions can be made during security risk analysis and planning to take into account dependencies and relationships between the security risks”.

Mackie goes on to argue that we often make subjective decisions about those factors that we focus on within a causal complex. The term ‘causal field’ refers to those factors that an investigator considers relevant to a particular investigation. If a cause does not appear within this subjective frame of reference then it is unlikely that it will be identified. This philosophical work has empirical support from the findings of West-Brown et al’s (2003) study into the performance and composition of Computer Security Incident Response teams. They describe the difficulties of ensuring that organizations and individuals broaden their view of the causal field to identify the different vulnerabilities that are exposed in the aftermath of security incidents. The problems of determining alternate causal fields are exacerbated by a number of factors identified by Meissner and Kassin (2002). They show that rather than improving accuracy in detecting deceit, training and prior experience make individuals more likely to identify ‘deceit’ rather than ‘truth’ in laboratory conditions. In other words, investigators cannot easily be trained to accurately identify whether evidence about the causes of an incident is true or not. Previous experience simply increases the likelihood that they will doubt the veracity of the information they obtain.

**Figure 1: Causal Fields and Primary Security Violations**

Figure 1 provides an overview of Mackie’s ideas. The causal field in this case concentrates on violations A, B and C. The term ‘violation’ refers to any act or omission that contravenes security requirements within an organization. Within the causal field, we can focus on particular issues that we raise to the status of ‘probable causes’. This is illustrated by the magnifying glass. For example, an investigator might be predisposed to look at the relationship between front office traders and back-office settlement staff. This
would be illustrated by the focus on the potential primary violation C in Figure 1. However, the causal field may not encompass a sufficient set of conditions and in this case Primary violation D is not within the range of issues being considered by the investigator. For instance, if the investigation focuses on the manner in which a rogue trader exploited vulnerabilities in reporting systems then correspondingly less attention may be paid to the role of other team members in detecting potential losses.

It is important to emphasize that this broader view of causation does not absolve individuals from responsibility for their role in security incidents. It is, however, important to recognize the diversity of other features within the causal complex of security incidents. In particular, the opportunities for individual violations are typically created by organizational and managerial problems. Individual criminal acts often form part of a more complex series of causes that are collectively sufficient for an incident to occur (Reason, 1997). In other words, many failures stem from ‘second order’ vulnerabilities. These describe problems that do not directly cause an adverse event but can help to create the conditions in which a security incident is more likely to occur.

Figure 2: Causal Fields and Secondary Security Vulnerabilities

Figure 2 provides an overview of secondary security violations. As can be seen, these problems contribute to primary failures. As we shall see, a lack of oversight in the separation between front office traders and the back-office staff responsible for settling accounts, represented by secondary failure 2, can create the vulnerabilities that are exploited by a rogue trader. This is illustrated by primary violation B in Figure 2. Alternatively, internal inspections by compliance teams following the model recommended by the Bank of England after the Baring collapse might help to detect such secondary vulnerabilities before they can be exploited. The successful barrier to secondary violation 1 in Figure 2 would illustrate this. An important aim of this paper is to extend the causal field of security investigations to consider these secondary causes of adverse events. This is illustrated in Figure 2 by moving the magnifying glass to the left. The dotted ellipse used to denote the causal field in Figure 1 could also be redrawn to show the extended scope of an investigation in this figure. Our emphasis on secondary violations is intended to guide the composition of a causal field, which Mackie argues can be a subjective and arbitrary process. These underlying secondary organizational, managerial and regulatory issues are an increasingly common factor in the assorted lists of ‘contributory factors’ that appear in security incident reports. We would, therefore, argue that these secondary violations deserve greater and more sustained attention.

To summarize, first order security violations lead directly to an incident. They are cited as the probable cause when, for instance, an individual attempts to place an unauthorized transaction. In contrast, secondary security vulnerabilities make these primary actions more likely. For example, inadequate management supervision can increase a rogue trader’s perception that their actions will not be detected. The increasing prominence of these secondary factors in regulatory reports suggests that more attention should be played to their role in the causal fields that guide security investigations.
THE CAUSAL ANALYSIS OF SECURITY INCIDENTS

It is clearly important that we learn as much as possible from those incidents that do take place if we are to reduce the likelihood and mitigate the consequences of security violations. A number of different tools and techniques can be used to support the analysis of these incidents. For instance, Julisch (2003) summarizes research into automated intrusion detection. He argues that over 90% of all alarms can be attributed to just over a dozen root causes. In consequence, rather than responding to individual alarms, investigators should focus on these more generic root causes using clustering methods that support the human analyst in identifying the underlying factors behind these warnings. Although this approach provides means of automatically clustering certain aspects of previous incidents, it cannot easily be applied to identify patterns in the organizational and managerial precursors to adverse events. In particular, it can be difficult to identify appropriate ways for representing and reasoning about these factors in security related incidents. Stephenson’s (2003) recent work on Colored Petri Nets for the analysis of ‘digital evidence’ avoids some of these limitations. He assessed the impact of the SQLSlammer worm on a multinational company. He was able to work back from the technical properties of the attack to identify the company’s business processes that made them vulnerable to this security threat. The formal Petri Net notation provided a common language for representing and reasoning about these different levels of analysis and hence could be used to move from the specifics of this incident to more general root causes. However, this work is based on a modeling language that was originally developed to support the design of concurrent systems. In consequence, it provides little direct support for the identification of root causes and contributory factors. The use of this approach is almost entirely dependent on the skill and expertise of the analyst. The lack of any supporting analytical methodology for the analysis of security incidents also makes it likely that two investigators will reach very different conclusions about the causes of an individual incident. This can help to identify a range of issues in the aftermath of an adverse event. Such inconsistency can also help to undermine the conclusions and recommendations that are drawn from an investigation.
Kilcrece et al’s (2003) work on organizational structures for security response teams reinforces the comments of the previous paragraph. It also highlights the consequences of the lack of methodological support for investigatory agencies. They argue “different members of the security team may conduct very different types of analysis, since there is no standard methodology”. In consequence, it is likely that effort will be duplicated both within response teams and across organizations as they address similar types of incidents. The lack of coordination and agreed procedures for the dissemination of root cause analysis makes it likely that similar patterns of failure will not be detected. This suggests that vulnerabilities will persist even though individual violations are identified. Without sharing this causal and contextual information, Kilcrece et al argue that the longer term recovery process will take longer and cost more, “problems that could have been prevented will instead spread across the enterprise, causing more down time, loss of productivity, and damage to the infrastructure”.

The US Department of Energy has recognized the importance of adopting appropriate methodologies for the root cause analysis of security incidents, particularly involving nuclear installations. OE Order 470.1 requires that this form of analysis be conducted and documented as part of any process “to correct safeguards and security problems found by Department of Energy’s oversight activities” (Jones, 2000). The intention is to ensure that any vulnerabilities are corrected in an ‘economic’ and ‘efficient’ manner. These methods are documented in the Department of Energy’s (2003) standard DOE-STD-1171-2003, the Safeguards and Security Functional Area Standard for DOE Defense Nuclear Facilities Technical Personnel. This requires that security personnel must demonstrate a working knowledge of root cause analysis techniques that can be applied to ‘determine the potential cause of problems’. They must be able to explain the application of root cause analysis techniques. In particular, they must be familiar with a number of specific approaches including causal factor analysis, change analysis, barrier analysis as well as management oversight and risk tree analysis. More detailed technical coverage of the application of these approaches is provided by the DOE (1992) standard DOE-NE-STD-1004-92, Guidelines for Root Cause Analysis. The adoption of root cause analysis does not, however, provide a panacea. A recent US General Accounting Office (GAO) report observed, “despite their importance, these assessments and analyses have not always been conducted”. The GAO argued that steps must be taken to ensure that Department of Energy staff and sub-contractors follow the recommended root cause analysis techniques in the aftermath of security incidents (Jones, 2000). In particular, it is important that staff be provided with sufficient training and case studies to enable them to apply techniques such as those described in the standard 1004-92.

The work of the US Department of Energy in the development of root cause analysis techniques has not been mirrored by similar developments in commercial and financial organizations. Recent interest in causal analysis from security consultancies, such as Price Waterhouse Coopers, and by regulatory organizations, including the Bank of England, has not led to any consensus about how such analysis should be performed. There is, therefore, a need to identify appropriate methodologies to probe beyond specific violations to identify the underlying ‘secondary’ vulnerabilities that create the context for most security incidents. It is for this reason that the following paragraphs present a case study in the application of root cause analysis techniques to a large-scale fraud investigation. The aim is to determine whether the tools and methods that have been developed by the US Department of Energy for investigations into nuclear security incidents might be more widely applied within the commercial sector. Later sections will motivate the decision to use these particular techniques. For now it is sufficient to observe that accident and incident analysis within the field of safety-critical systems have been supported by a vast range causal investigation tools. Many of these are summarized in Johnson (2003). In contrast, we have chosen to focus on those approved by the US DOE because these techniques are well documented and have at least a limited track-record within the limited field of nuclear security investigations.

OVERVIEW OF THE ALLFIRST CURRENCY TRADING LOSSES

The remainder of this paper is illustrated by a case study involving the loss of approximately $750 million in currency transactions from Allfirst, a subsidiary of Allied Irish Bank. This case study is appropriate because it illustrates how managerial difficulties, human ‘error’ and technical security failures combined to create systems weaknesses. The account used in this paper draws heavily on the report to AIB by the
Promontory Financial Group and by Wachtell, Lipton, Rosen and Katz (Promontory, 2002). Other sources have also been used and these are acknowledged at the point at which their material is introduced.

In 1983, the Allied Irish Bank (AIB) acquired a stake in Allfirst, then known as the First Maryland Bancorp. This stake grew until by 1989, AIB had taken acquired First Maryland through a merger. AIB planned to diversify its operations in North America. They believed that this could best be achieved by allowing Allfirst a large amount of local autonomy. Allfirst continued have its own management team and board of directors. However, stronger control was retained over Treasury operations via the appointment of a senior AIB executive to oversee these operations. Prior to his appointment in 1989, there had only been a minimal history of currency trading at Allfirst with limited risks and a limited budget. In 1990, however, a trader was recruited to run proprietary trading. These operations continued relatively successfully until the first incumbent of this post had to be replaced in 1993. John Rusnak was recruited from a rival bank in New York, where he had traded currency options since 1989. One aspect of his recruitment was the desire by Allfirst to exploit a form of arbitrage that Rusnak specialized in. This took advantage of the differences in price between currency options and currency forwards. In simple terms, an option is an agreement that gives the buyer the right but not the obligation to buy or sell a currency at a specified price on or before a specific future date. If it is exercised, the seller must deliver the currency at the specified price. A forward is a contract to provide foreign exchange with a maturity of over 2 business days from the transaction date.

Rusnak’s activities can be seen in terms of the primary violations described in Figure 1. He created bogus options to hide losses that he had sustained in currency trading. These catalytic events exploited underlying vulnerabilities, similar to those sketched in Figure 2. For example, the immediate report into the fraud identified ‘numerous deficiencies’ in the control structures at Allfirst. In line with Mackay’s assertions about causal complexes, the report went on to argue that ‘no single deficiency can be said to have caused the entire loss’ (Promontory, 2002). The underlying vulnerabilities included the failure of the back-office to confirm Rusnak’s bogus options with the counterparties involved in the transaction. Such checks might have revealed that these counterparties had no knowledge of the fictitious transactions that Rusnak said they were involved in.

Many of the secondary problems at Allfirst relate to their organizational structure. Allfirst’s treasury operations were divided into three areas. Rusnak’s currency trading was part of the front office. The middle office was responsible for liability and risk management. The back-office was responsible for confirming, settling and accounting for foreign exchange and interest rate derivatives trades, including those initiated by Rusnak. Allfirst espoused the policy of having the back-office confirm all trades, following industry practice. The initial reports speculate that Rusnak may have put pressure on his colleagues not to confirm all of his options trades. Figure 4 sketches the relationship between the different reporting structures in the Allfirst treasury. Rusnak formed part of a relatively small and specialized group in the Foreign Exchange area. This diagram also illustrates some of the potential vulnerabilities in the reporting mechanisms within the bank. The Allfirst Treasurer was responsible both for ensuring profitable trading and for ensuring effective controls on that trading. Subsequent investigations also revealed concerns about the Treasury Funds Manager’s position. Not only did they direct many of the Treasury operations but they also controlled many of the reporting procedures that were used to monitor operational risks. The Vice President for Risk Control, therefore, devised a plan so that asset and liability management reports as well as risk control summaries would be directed to senior management through his office. Unfortunately, this plan does not seem to have been implemented before the fraud was detected.
The previous paragraphs have summarized the primary violations and secondary vulnerabilities that contributed to the Allfirst fraud. The failure to investigate potential security issues once they had been identified also illustrates tertiary failures of the type described in the opening sections of this paper. The main aim behind this overview has been to provide a concrete example of the complexity of causal arguments in security incidents. The following sections use this initial analysis to illustrate how root cause analysis techniques can be extended from accident investigations to examine a wider class of security failures.

**INTRODUCTION TO ROOT CAUSE ANALYSIS TECHNIQUES**

Root cause analysis techniques provide tools for identifying the elements of a causal field from a mass of other contextual factors. In Mackay’s terms they can also be used to determine the composition of various causal complexes within such a field of relevant factors. Recall that each causal field is one of several possible combinations of factors that might lead to an adverse outcome. Each individual factor within a field is necessary but, typically, not sufficient for an incident to occur. As previous sections have argued, without appropriate tools it is likely that analysts will miss important factors within one or more of these causal fields. It is also likely that individual differences will lead to inconsistency between the findings of multiple independent investigators. In other words, there are likely to be significant differences over whether or not a particular factor is a necessary cause of an adverse event. This can be illustrated by the subsequent debate and litigation as to whether the prime brokerage accounts played a significant role in the causes of Allfirst’s eventual loss. Root cause analysis techniques provide tools and techniques that can be used to encourage agreement over those factors, violations and vulnerabilities, that contribute to a security failure.

**Barrier Analysis**

The previous summary of the Allfirst fraud provides a false impression of the problems that face investigators in the aftermath of a security violation. The outcome is often, but not always, fully understood. Far less is known about the vulnerabilities that created the context for particular violations. In consequence, most investigations begin with a prolonged period of elicitation where evidence is gradually gathered about the course of an incident. Barrier analysis can be used to support these parallel activities. It also provides documentary evidence to help demonstrate that investigators have considered a broad range of causal fields.
Barrier analysis is based on the idea that most security-critical systems rely on counter-measures or barriers that are intended to prevent a security hazard from gaining access to or adversely affecting a target. Figure 5 provides an overview of the central ideas in Barrier Analysis. As can be seen, a security hazard must pass through a series of potential barriers before they can reach the ultimate target. The weaknesses in these various barriers can be seen as the vulnerabilities mentioned in previous sections. The events that undermine these barriers have been called violations. In Figure 5, the final barrier denies access to, or prevents the security hazard from affecting, the target. This typifies the way in which a final layer of defenses can make the difference between an unsuccessful attack and a security breach. In such circumstances, incident investigations provide important insights both about those barriers that failed and those that acted to protect the target from a security hazard.

**Figure 5: Targets, Hazards and Barriers**

Barrier analysis is based on the idea that most security-critical systems rely on counter-measures or barriers that are intended to prevent a security hazard from gaining access to or adversely affecting a target. Figure 5 provides an overview of the central ideas in Barrier Analysis. As can be seen, a security hazard must pass through a series of potential barriers before they can reach the ultimate target. The weaknesses in these various barriers can be seen as the vulnerabilities mentioned in previous sections. The events that undermine these barriers have been called violations. In Figure 5, the final barrier denies access to, or prevents the security hazard from affecting, the target. This typifies the way in which a final layer of defenses can make the difference between an unsuccessful attack and a security breach. In such circumstances, incident investigations provide important insights both about those barriers that failed and those that acted to protect the target from a security hazard.

<table>
<thead>
<tr>
<th>What?</th>
<th>Rationale</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hazard</td>
<td>Currency trading losses concealed by fraudulent use of the Bank’s assets.</td>
</tr>
<tr>
<td>Targets</td>
<td>Allfirst’s risk exposure and ultimately the Bank’s assets.</td>
</tr>
</tbody>
</table>

Table 1: Hazard and Target Identification

Table 1 illustrates the initial stages of a barrier analysis. Investigators must first identify the hazard and targets involved in a security incident. During these initial stages, the analysis is conducted at a relatively high level of abstraction. The investigation progresses by examining the barriers that might prevent a hazard from affecting the targets. Analysts must account for the reasons why each barrier actually did or might have failed to protect the target. Table 2 illustrates the output from this more detailed stage of analysis. As can be seen, the barriers are those defenses that were intended to prevent undisclosed currency trading losses from distorting the bank’s risk exposure and reducing the Bank’s assets. As can be seen, the Value at Risk calculation (VaR) was one of the barriers that were intended to protect Allfirst’s risk exposure. The policy of ensuring that the back-office confirmed all transactions should also have kept bogus trades off the balance sheet. Similarly, the foreign exchange rates used by the middle and back-office should have come from an independent source to reduce the risk of manipulations masking the true risk of any options. Analysis progresses by identifying the reasons why each of these barriers was breached. In other words, this approach helps to identify the vulnerabilities that might have been exploited in the course of any security incidents. The development of such a table also helps to focus the gathering of evidence. Investigators must obtain proof that supports the assertions made in Table 2. For example, the Value at Risk calculation provided Allfirst with a statistical measure for the ‘worst case’ losses that might be expected from a given portfolio. Monte Carlo techniques were used on a range of hypothetical spot and volatility rates for different currencies. The VaR was derived from the tenth worst outcome obtained from the stochastic model. However, as can be seen in Table 2, Rusnak distorted the VaR calculations by creating bogus options that hedged, or mitigated, the risks associated with his real trades. Similarly, the VaR calculation was distorted by the manner in which Rusnak persuaded an employee in the...
Risk Control Group to use data from his PC to calculate the extent of ‘holdover transactions’. Rusnak often held very large ‘open’ positions. In general terms, an open position is a deal that is not yet reversed or settled. The investor is subject to exchange rate movements. Hence, the larger the open position then the greater the risk denoted by the VaR. Any transactions that could not be fully processed at the end of the day were ‘heldover’ until the following morning. These were not entered into the bank’s trading software and so were not included in the VaR calculations. It would have been relatively easy to check whether trades were being hidden. This could be done by adding them to the ‘holdovers’ and then removing them the next day. For example, internal audit might have compared the list of holdovers against the next day’s trading activity to ensure that the trades had, in fact, been entered into the trading software. The abuse of the holdover system was so extreme that several transactions were left in the system for three consecutive days without being entered into the trading software. The manner in which Rusnak was able to enter false transactions into the DEVON system represents a further way in which the VaR barrier was undermined. This software was used to record all transactions on the prime brokerage accounts. In particular, he entered false forward transactions and then reversed them prior to the next settlement date. This enabled him to create transactions that would alter the VaR, for instance by hedging other transactions. These bogus trades could then be removed from the system before any settlement took place.

<table>
<thead>
<tr>
<th>Barrier</th>
<th>Reason for failure?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Value at Risk Calculations</td>
<td>Rusnak created bogus options to hedge his real positions and hence those options offset or reduced the Value at Risk (VaR) from his activities.</td>
</tr>
<tr>
<td></td>
<td>Rusnak provided false figures for ‘holdover’ transactions. An employee in the Allfirst risk control group used a spreadsheet to calculate the VaR that relied on files obtained from Rusnak’s PC.</td>
</tr>
<tr>
<td></td>
<td>Rusnak was able to enter false foreign currency exchange forward transactions in the DEVON system and reverse them before the settlement date, manipulating the transactions that were used to calculate the VaR.</td>
</tr>
<tr>
<td>Back-office confirmation of bogus options with counterparties.</td>
<td>Back-office staff did not confirm all of Rusnak’s transactions. Promontory group report refers to bullying and cajoling.</td>
</tr>
<tr>
<td></td>
<td>Difficulty of confirming Asian trades in middle of US night persuades treasury senior staff to decide not to confirm offsetting pairs of trades with these markets.</td>
</tr>
<tr>
<td>Foreign exchange rate confirmation from independent sources.</td>
<td>Rusnak opposes direct drop of Reuters prices into middle and back-office systems.</td>
</tr>
<tr>
<td></td>
<td>Risk assessment analyst fails to communicate concerns to senior management.</td>
</tr>
<tr>
<td></td>
<td>Decision not to invest in dedicated Reuters terminals for middle and back-office staff.</td>
</tr>
<tr>
<td></td>
<td>Delays in implementing back-office ‘computer disk’ to store prices direct from Reuters once senior management alerted to downgrade in second quarter 2001 risk assessment report.</td>
</tr>
</tbody>
</table>

Table 2: More Detailed Barrier Analysis

Table 2 identifies a number of other barriers that were intended to protect Allfirst’s risk exposure from bogus transactions. In particular, it is a common assumption that back-office staff will confirm that all trades have actually been made or placed with counterparties. Rusnak was able to create bogus hedging positions because this did not happen at Allfirst. The subsequent investigations revealed that Rusnak had pressurized responsible staff in the Allfirst treasury so that they did not routinely check all of his transactions. In addition to such persuasion, a number of other reasons explain this apparent violation of company policy. Firstly, Rusnak typically created pairs of offsetting transactions so it looked as though the risk in one trade was matched by hedging positions in another. The lack of any cash payments and the problems associated with contacting traders in the Asian markets may have combined to reassure the staff member involved that confirmation was not always necessary.
A final barrier mentioned in Table 2 is the practice of obtaining currency exchange rate information from an independent source. In the past this had been done by printing the exchange rates directly from the Reuters feed to the front office and then cross checking these values against the rates used by the middle and back-offices. The treasury risk control analyst argues that this was error prone and cumbersome and that a spreadsheet might be developed to speed access across these departments. The intention was that the software would directly interface with the Reuters link. However, this potential barrier to exchange manipulation was undermined by Rusnak’s insistence that he have access to the spreadsheet so that he could use the exchange rates to monitor his VaR. The treasury risk control analyst, therefore, allowed an architecture to develop whereby the Reuters feed to the front office was downloaded onto Rusnak’s PC. The middle and back-office system then drew data from these files on Rusnak’s machine. The risk assessment analyst subsequently met with the treasury risk control analyst and queried whether this was appropriate. Table 2 also captures the observation that ‘the risk assessment analyst failed to communicate concerns to senior management’. Senior management did eventually become aware of this procedure when the risk assessment group downgraded the first quarter 2001 risk assessment report from Good to Weak. However, the Barrier Analysis also records that there was a significant delay before the back-office was equipped with their feed and disk to store the Reuters currency information.

As can be seen, the barrier analysis represented in Figure 5 encourages analysts to consider both the underlying vulnerabilities and violations that combine to compromise the security of many complex systems. For instance, Rusnak’s manipulation of the ‘holdover’ transactions was only possible because there was an underlying vulnerability created by the failure to check that such trades had actually been entered during the next working day. Similarly, Rusnak’s manipulation of the Reuter’s feed was only possible because of the decision not to provide middle and back-office staff with their own dedicated links.

Change Analysis

Change analysis provides a similar form of support to that offered by barrier analysis. Rather than focusing on those defenses that either protected or failed to protect a potential target, change analysis looks at the differences that occur between the actual events leading to a security incident and ‘ideal’ operating procedures. For example, the actual mechanisms used to obtain pricing information might be compared with those described in a company’s risk control documents. Table 3 provides an example of change analysis. The first column describes the ideal condition. In some applications of this technique, the first column is instead used to represent the practices and procedures that held immediately prior to a security incident. This is an important distinction because the causes of an adverse event may have stemmed from inappropriate practices that continued for many months. In such circumstances, the change analysis would focus less on the conditions immediately before the incident and more on the reasons why practice changed from the ideal some time before the mishap.

Table 3 shows that Rusnak’s supervisors should have examined his positions and trades in greater depth given the overall size of his positions. This ‘normative’ statement can be justified by referring to a range of Allfirst and AIB documentation on internal audit and risk control (Promontory, 2002). The middle column indicates several of the ways in which Allfirst practice differed from this norm. No one noticed that many of Rusnak’s options expired unexercised on the day that they were created. This enabled him to leave bogus balancing transactions on the book. The longer-term bogus transactions avoided suspicion because they appeared to be hedged by the short-term options that expired unexercised. Normal security precautions such as telephone tapping and logging were not used. This deprived risk control managers of important sources of information that might have alerted them to the lack of communication with the counterparties on many of the bogus options. Finally the lack of scrutiny on Rusnak’s positions is revealed by the failure to reconcile his daily profit and loss figures with the general ledger at Allfirst. One consequence of this was that Rusnak was able to develop trades well beyond his daily limits, for example by the abuse of the holdover system mentioned in previous sections.
Table 3: Change Analysis

Table 3 also illustrates the argument that normal auditing practice should ensure that suggestions made by audit, risk assessment and supervisory examinations are followed through until they are either implemented or reasons for their rejection are adequately documented. In contrast, several recommendations were ignored or only implemented in a piecemeal fashion during the Allfirst fraud. The lack of systematic monitoring for auditing recommendations created opportunities for Rusnak. The resulting vulnerabilities included a considerable delay in establishing independent sources for currency pricing information. This enabled Rusnak to manipulate the VaR calculations for his trading activities.

An important benefit of change analysis is that the ‘ideal’ conditions in these tables can be used to identify recommendations. This is not straightforward. For instance, stating that staff and management should follow the company’s risk control procedures does not provide any guarantee of compliance. The prior/ideal condition column in the change analysis tables can, however, provide a starting point for the identification of more detailed recommendations. In Table 3, investigators might argue that a monitoring system should be introduced to trace the implementation of audit, risk assessment and supervisory examinations. It should then be possible for senior management to use the system to ensure the implementation of necessary interventions recommended by these internal audits. Had such a system been adequately implemented then Allfirst might have avoided or minimized the delays associated with the development of an independent currency pricing system from the middle and back-offices.

A number of limitations restrict the utility of change analysis. For instance, they often introduce a form of hindsight bias. Norms were not followed because violations were able to exploit existing vulnerabilities. It is, therefore, tempting to argue that existing rules and regulations should be applied more diligently in the future. This is a dangerous argument. It assumes that existing procedures and practices were sufficient to ensure the security of a system. Further limitations affect both Barrier Analysis and Change Analysis. These techniques can be used to structure the initial analysis of a security incident. They guide investigators by providing a framework of important concepts as they gather information about what should
have happened and what actually did occur during particular violations. They do not, however, provide more detailed support for the modeling that is often necessary in order to understand the complex manner in which different events and causal factors combine over the course of a security incident. Event based modeling techniques can be used to avoid this limitation during the reconstruction of complex failures.

VIOLATION AND VULNERABILITY ANALYSIS (V² ANALYSIS)

Many different event-based techniques have been developed to support the root cause analysis of safety-related incidents. These include Events and Causal Factors charting (ECF), Multilinear Events Sequencing (MES) and Sequential Timed Event Plotting (STEP). Brevity prevents a detailed analysis of each of these approaches; the interested reader is directed to Johnson (2003). The key point is that several of these techniques have also been used to analyse the underlying vulnerabilities and specific violations that lead to security related incidents (US Department of Energy, 1992). Most previous applications have been within the specific context of nuclear energy and weapons development. A further limitation to the more general application of these techniques is that they provide little specific support for the analysis of security incidents. Hence, the basic components in these event-based techniques are unchanged from their use in safety-related applications even though the details surrounding these ‘dependability’ failures can be very different.

In contrast, Figure 6 provides an example of Violation and Vulnerability (V²) analysis. This extends an event based modelling technique to deliberately support the identification of root causes for a wide range of security related incidents. The underlying approach is similar to the existing ECF, MES and STEP techniques, mentioned above. This V² diagram is constructed around a number of events that are denoted by rectangles. For example, ‘AIB insert senior manager as Allfirst treasurer’ and ‘Treasurer is appointed to key AIB group marketing strategy committee’ are both shown as events in Figure 6. These are made more likely by a number of contributory factors that are shown by ellipses. For instance, the decision to insert one of the AIB executives as the Allfirst Treasurer led to a situation in which some viewed the treasurer as a form of ‘home office spy’. This contributed to the exclusion of the formed AIB executive from some senior management decisions at Allfirst.

Figure 6 focuses more on the contextual factors than on specific events during the Allfirst fraud. It also maps out a range of conditions that formed the background to the more detailed events that are mentioned in previous sections. This is deliberate because an important objective behind the use of this modeling technique is to trace the roots of a security violation back into the underlying vulnerabilities within the operations of a company, such as Allfirst. Vulnerabilities can be thought of as a particular type of contributory factor. As mentioned in Figures 1 and 2, they create the opportunity for the violations that occur during security incidents. In Figure 6, vulnerabilities relate to the dual reporting structure between AIB and Allfirst. They weakened the supervision of the Treasurer’s activities in the lead-up to the fraud. This vulnerability is denoted by the double ellipse at the bottom right of figure 6. Subsequent V² diagrams can be used to map out the precise manner in which this particular contributory factor acted as a precondition for Rusnak’s violations.
Figure 6: A V² Diagram of the Background to the Allfirst Fraud
Figure 6 illustrates the way in which $V^2$ diagrams can be used to look beyond the particular violations that lead to a fraud. This is important if investigations are to accurately identify the underlying managerial and organizational factors that might lead to future security problems. For instance, one response to the events at Allfirst would simply have been to focus legal retribution on the trader. This would, however, have ignored underlying problems in the relationship between AIB and Allfirst, including the supervision of key Treasury staff. This point is made more forcefully in the recommendations that emerged in the immediate aftermath of the fraud: ‘In light of the foregoing considerations, AIB should consider terminating all proprietary trading activities at Allfirst, and all customer trading activities at Allfirst should be relocated to the AIB branch in New York. While the salespeople may continue to be located in Baltimore, any price-making and trade execution should be done in New York, under the direct supervision of AIB treasury’ (Promontory, 2002).

Figure 7 continues the Violations and Vulnerability analysis by documenting the events leading to the hiring of Rusnak by Allfirst. Prior to 1989, Allfirst had only engaged in limited currency trading. This contributed to the decision to recruit a specialist to run their proprietary trading business. During this period, trading was focused on directional trading, in other words profits were dependent on forecasting the future price of a currency as it moved up or down on the markets. The senior trader left Allfirst and a further event in Figure 7 is used to show that the ‘Treasury funds manager heads the search for a new trader’. This leads to an offer being made to Rusnak. The decision to make this offer was supported by recommendations from his previous employers at Chemical Bank. His appointment was also supported by the Allfirst Senior Management’s interest in Rusnak’s non-directional trading. This will be described in more detail in subsequent $V^2$ diagrams. Figure 7 also illustrates how these various events, together with a number of additional contributory factors lead to a further security vulnerability. Allfirst’s efficiency committee suggested that the treasurer scale-back proprietary currency trading. However, the senior management interest in Rusnak’s non-directional approach helped to focus the cutbacks in more conventional forms of currency trading. The senior management interest also created a situation in which the Treasury funds manager was highly protective of Rusnak and his activities. These various factors combined to weaken the monitoring and reporting procedures that were established to control the risks associated with his activities. When Rusnak’s immediate trading manager resigned, his post was not filled. Lack of funds prevented a renewed appointment and so Rusnak now reported directly to the treasury funds manager who, as we have already seen, was protective of his non-directional trading strategies.

Analysts can use $V^2$ diagrams to map out the mass of contextual details that emerge during an investigation. Change and Barrier analysis can be used to identify these contributory factors and events. A number of other approaches, such as Conclusion, Analysis and Evidence diagrams and Why-Because Analysis, have been developed within the field of accident analysis to exploit more narrow definitions of causal relationships than those illustrated in Figure 7 (Johnson, 2003). Alternatively, Multilinear Event Sequencing (MES) is one of several techniques impose additional formatting constraints on diagrams that are similar to those shown in this paper (US Department of Energy, 1992). MES uses a grid in which the events relating to particular actors or agents had to be shown along the same row. Columns were then used to denote the flow of events over time. Each event had to be shown to the right of the events that occurred before it. In contrast, $V^2$ diagrams take a more relaxed approach. It can be difficult to establish the exact timing for many events. This problem can be even worse for contributory factors. For instance, when should an investigator show that ‘Senior Management were intrigued by Rusnak’s non-directional trading approach’? This sentiment seems to have emerged over a prolonged period of time and cannot easily be associated with particular meetings or events, especially in the aftermath of a security incident. Similarly, other events affect many different actors in an adverse event. In Figure 6, several different managers supported the appointment of the Treasurer on the AIB and Allfirst committees. These events would have to be widely distributed across many different columns in a MES diagram adding to the complexity of constructing and maintaining these representations. It is for this reason that $V^2$ diagrams relax some of the constraints that guide Multilinear Event Sequencing. Arrows represent relationships or constraints. They do not represent necessary causal relationships. For example, the protective attitude of the Treasury Funds Manager did not ‘cause’ the flaws that affected the reporting and monitoring of Rusnak’s work. The fraud may even have occurred if the Treasury Funds Manager had not been so protective. However, the manner in which he shielded the trader from subsequent enquiries did have a profound impact on the underlying vulnerability illustrated in Figure 7.
Figure 7: A V2 Diagram of the Events Leading to Rusnak’s Appointment and Flaws in his Reporting Structure
Figure 8 extends the $V^2$ analysis towards the events surrounding Rusnak’s fraudulent activities. As can be seen, he initially created the impression that he specialized in a form of arbitrage by taking a profit from differences in the exchange rates between different markets. In particular, he claimed to make profits by holding a large number of options that were hedged by balancing positions in the cash market. These observations are denoted in Figure 8 by the contributory factors at the top-right of the diagram. The contributory factors at the top-left show that most of his trades were simpler than many at Allfirst had supposed. They involved linear trades based simply on predicted fluctuations in currency rates. This led him to buy significant quantities of Yen for future delivery. The subsequent decline in value of this currency prior to delivery left Rusnak with a loss. Combined with the image that he had fashioned for his trading activities, the loss may have created a situation in which he felt under pressure to hide the outcomes from his options on the Yen. This analysis of the top components in Figure 8 raises a number of important issues about the construction of $V^2$ diagrams. It can be argued that Rusnak’s creation of a false impression about the nature of his trades should be ‘promoted’ from a contributory factor to either a violation, and therefore be linked to specific events, or vulnerability. The tension between his claimed trading techniques and his actual methods helps to explain many of his subsequent actions. It can equally well be argued that such tensions are widespread within many financial organizations. Several studies have pointed to the psychological characteristics and personality attributes of successful traders (Tvede, 1999). It has been argued, for instance in Oberlecher’s (2004) study of the psychology of foreign exchange markets, that the same attributes that create these tensions between action and appearance may also be important ingredients in the makeup of successful traders. The meta-level point here is that $V^2$ analysis forces investigators to consider whether or not each contributory factor could be considered a potential vulnerability and also whether each event in the context of a security incident might also be labeled a violation. There is no automatic or algorithmic process to support this analysis.

Figure 8 also illustrates the mechanisms that Rusnak used to hide his losses from directional trading on the Yen. These have been briefly outlined in previous sections. Initially, he began by creating a bogus ‘deep in the money’ option. Recall that such an option has a price that is significantly below the current spot-price and hence it is high risk for the vendor. Such options attract high premiums, especially if they can be exercised in the short term when the spot price is unlikely to fall below the level of the quoted option. Allfirst, therefore, had a significant potential liability. At the same time, he created a second balancing bogus option with the same counterparty. This is represented in Figure 8 by the violation labeled ‘Rusnak creates balancing option as if Allfirst have paid a large premium to buy currency weeks later involving the same counterparty’. This made it look like Allfirst’s original liability was offset by the asset value of the second option. Allfirst should have paid a correspondingly large premium to obtain this second option even though no cash would actually have changed hands because the two premiums balanced each other and were drawn against the same parties. The crucial difference between these options was that the first one, representing Allfirst’s liability, was set up to expire within 24 hours. The second, representing Allfirst’s fictitious asset, expired several weeks later. Rusnak knew that neither option would ever be exercised because they were bogus deals. However, for the period between the expiry on the first option and the end of the second, he was able to create the appearance of a genuine asset on the Allfirst books. This could be used to offset his own genuine losses.

These deals made no sense for a number of reasons. Firstly, the risk exposure on each of the options was quite different given that one expired in 24 hours while the second typically lasted for several weeks. In such circumstances, the options should have attracted very different premiums and so were unlikely to balance each other out. Secondly, the ‘deep in the money’ options involved in the first bogus trade should have been exercised by the counterparty. A series of similar options failing to be acted upon should have alerted company management to potential fraud. However, as Figure 8 also shows, Allfirst managers did not have access to a list of those options that had expired without being exercised within 24 hours of them being placed. This is denoted by the vulnerability on the left hand side of the $V^2$ diagram. Prior to September 1998, Rusnak covered his tracks by creating bogus confirmations from the supposed counterparties to these transactions. The confirmations were intended to provide evidence that both parties had agreed upon these trade options. After that time, Rusnak managed to persuade the back-office staff not to pursue these confirmations for his trading activities. As can be seen from the $V^2$ diagram, their failure to confirm the transactions is partly explained by the difficulty of establishing contact with many of Rusnak’s brokers who worked in the Asian offices of the counterparties. The trader’s office hours often created considerable communications difficulties for Allfirst’s back-office staff. Figure 8 also uses a triangle continuation symbol, labeled with a ‘2’, to carry the analysis from the events surrounding Rusnak’s appointment to the start of his fraud. As can be seen, flaws in the reporting and monitoring procedures for Rusnak’s activities made it more likely that he would be able to persuade back-office staff not to confirm the matching pairs of bogus trades. These flaws stemmed in part from senior management’s desire to support his ‘novel’ forms of arbitrage.
Figure 8: A $V^2$ Diagram of Rusnak’s Initial Balanced-Options Fraud
Figure 8 also captures the cyclical nature of Rusnak’s fraud. Eventually the second option in each bogus pair would expire unexercised. At this point, the large and fictitious asset would disappear from Allfirst’s books. It was, therefore, important that Rusnak continue to generate these option pairs if his fraud was not to be discovered. This is indicated by the arrow from the bottom right of Figure 8 between ‘2nd option expires unexercised, original loss now needs to be covered again’ back to the contributory factor ‘Rusnak is under pressure to hide his losses’. In previous V² diagrams, rectangles have been used to denote specific events. In contrast, Figure 8 shows the structure of Rusnak’s fraud using generic events that represent a class of similar violations. It would, of course, be possible to construct a more specific model that represents each of the individual trades that made up this pattern within the security incident. However, the level of detail illustrated in the previous diagram is appropriate for most stages of an investigation. At this stage of the analysis, there is little to be gained from individually identifying the unwitting counterparties to Rusnak’s options trades.

Figure 9 continues the V² analysis of the Allfirst fraud. The ability to represent change over time is important because many security incidents develop over months or years. The individuals and groups involved often alter their behavior in response to external events and the audit mechanisms that are used to detect any continuing vulnerabilities. This diagram shows how Rusnak exploited further opportunities to expand both his trading activities and the range of bogus trades that were required to conceal his mounting losses. The top right event in Figure 9 denotes that Rusnak was offered net settlement agreements with a number of financial institutions (Promontory, 2002). These eventually developed into ‘prime brokerage accounts’. Such facilities enabled the broker to settle spot foreign exchange transactions with the counterparties. Each of these individual trades was then rolled together into as larger forward transaction between the broker and Allfirst that could be settled on a fixed date every month. As can be seen, these agreements simplified multiple transactions between Allfirst and the counterparties into a smaller number of larger transactions with the brokers. This simplification had two effects. Firstly it reduced the number of operations for the Allfirst back-office. Secondly, it made it difficult for the back-office and others within Allfirst from monitoring the individual trades that were being roller together within Rusnak’s prime brokerage accounts. This potential vulnerability is represented half way down Figure 9 on the right hand side.

The problems of monitoring transactions through the prime brokerage accounts together with the ability to roll together individual transactions for periodic settlement together combined to create a situation in which Rusnak could exceed the limits on his trading that were routinely insisted upon by Allfirst. His ability to increase the scope and scale of his trading is shown in Figure 9 to have increased the amounts of his loses in both forward and spot transactions. In order to cover his losses, another cycle emerged in which he generated more bogus transactions using the balancing options approach, described in previous sections. Rusnak was also able to exploit vulnerabilities in the DEVON software. This was used to track trades across the prime brokerage accounts. He was able to enter bogus transactions into the system and then reverse them before the monthly settlement period. As can be seen, however, Figure 9 does not provide sufficient details about the nature of the underlying problems with the DEVON application. The vulnerability symbol is annotated with the comment; ‘DEVON system vulnerabilities (further analysis?)’. The V² notation could be revised to explicitly represent this need for additional analysis. More symbols could be used to show those events and contextual factors, violations and vulnerabilities that have only been partially analyzed. This has not been done, however, in order to minimize the amount of investment that must be made in training to both read and eventually develop these diagrams.

The right-hand, lower portion of Figure 9 illustrates a series of events that threatened Rusnak’s activities. It began when the Allfirst treasurer decided to introduce a charge on those activities that used the bank’s balance sheet. Such a change would provide greater accountability, for example by exposing whether the profits generated by an activity actually justified the work created for those who must maintain the balance sheet. Questions began to be asked about whether the apparent profits from Rusnak’s activities could justify his use of the balance sheet. The total volume of currency traded had risen rapidly over the year to January 2001 but net trading income remained almost the same. A significant proportion of this rise can be attributed to Rusnak’s various trading activities. He was, therefore, told to reduce his use of the balance sheet. This not only curtailed his legitimate trading activities but also placed tight constraints on many of the bogus trades, even if many of those trades only made a fleeting appearance on the Allfrist books before being reversed. He had to identify an alternate source of funds to offset his previous losses and those that continued to accrue from his legitimate trading activities.
Figure 9: A V² Diagram of Rusnak’s Manipulation of Prime Brokerage Accounts
Figure 10 traces the Allfirst fraud from the point at which senior management began to question Rusnak’s use of the bank’s balance sheet. This is denoted by the continuation symbol, labeled 4, connecting this image with the V² diagram in Figure 9. Rusnak’s need to find an alternate source of funds led him to sell long-term options that were deep in the money. As mentioned previously, these options quoted a strike price that was far above the currency’s current spot price. Hence, the options represented a relatively high-risk for Allfirst and attracted a corresponding premium. However, Figure 10 also uses a contributory factor to denote that these ‘deep in the money options can be viewed as a form of loan’ and that ‘Rusnak would need to get these liabilities off the books’. Allfirst would have to redeem them when the options were redeemed. Figure 10 denotes a further violation as Rusnak created bogus transactions to indicate that the original options had been repurchased. These activities again involved Rusnak’s use of the balance sheet and so the Allfirst treasurer placed a limit of $150 million on his trades.

Previous V² diagrams have shown how Rusnak was able to manipulate the DEVON system to conceal some of his transactions via the prime brokerage accounts. Figure 10 shows some of the consequences of these manipulations through the continuation symbol, labeled 5, that links back to the previous diagram. The misuse of the DEVON system, combined with the ‘bogus’ repurchasing of ‘deep in the money’ options distorted the Value at Risk (VaR) calculations that were introduced in previous sections. Figure 10 also illustrates further ways in which this risk assessment tool was undermined. Rusnak used ‘holdover transactions’ to disguise some of his trades. These transactions usually occurred after it was possible for them to be included in the day’s accounts. They were, therefore, held over until they could be processed during the next trading day. Internal audit and risk control were aware that Rusnak was responsible for a large number of these transactions but they did not investigate. This observation is illustrated by the vulnerability at the top right of Figure 10. Holdover transactions were not entered directly onto the bank’s trading software. There were no checks to determine whether transactions were actually entered into the following day’s trading. All of these vulnerabilities can be seen as causal factors in a violation of audit procedures whereby Rusnak directly supplied risk group employees with on-line data for his holdover transactions.

The VaR was the main statistical measure used by the Allfirst senior management to monitor currency trading. Figure 10 also traces other mechanisms that Rusnak used to undermine reporting within the bank. The contributory factors on the right hand side of this diagram record how currency revaluations were initially printed from a Reuters feed and then were manually checked against the back-office records. Several audits had stressed the need to ensure that this data was derived from an independent source that could not easily be manipulated by individual traders. This potential vulnerability was addressed in a treasury control analyst’s proposal that Reuter’s data could automatically be fed into the bank’s internal systems. However, as Figure 10 illustrates, Rusnak was able to violate these security concerns by arguing that he needed constant access to changing currency rates across the bank in order for him to monitor and assess his VaR calculation. This combined with the associated expense of providing dedicated Reuter’s feeds into the back-office persuaded treasury to enable Rusnak to feed data from his own software into the rest of the Allfirst systems.

The previous paragraph illustrates how V² analysis provides a framework for the detailed investigation of security related incidents. Rusnak’s ability to persuade the treasury control analyst that it would be acceptable for him to pass on pricing information to other bank systems requires further analysis. The figure quoted for dedicated feeds and the apparent disregard of previous audits could be the focus for subsequent investigation using complementary techniques, including the Barrier and Change Analysis. The key point is, however, that these diagrams provide an overview of the complex events and contributory factors that lead to security incidents. The resulting sketches can be shown to other members of multi-skilled investigatory teams so that they are built up over time. Peer review can also help to ensure that the resulting analysis captures both the primary violations and secondary vulnerabilities that lead to adverse events.
Figure 10: A $V^2$ Diagram of Rusnak’s ‘Deep in the Money’ Options and the VaR Calculations
The analysis of the failed barriers to Rusnak’s fraud continues in the V^2 diagrams. Figure 11 also shows that one outcome of the risk assessment analyst’s decision to pursue the currency feed personally was that she asked Rusnak to email her a copy of his spreadsheet that was used to pass on values to the back-office. She ‘immediately discovered Yen and Euro values were corrupted’ and then downgraded the control market risk from good to weak’ and the ‘quality of risk management also falls to acceptable’. These actions finally acted as a trigger for a more senior involvement. However, by this time Rusnak had halted his price manipulation and so when back-office staff checked the values they tallied with the external sources.

Allfirst internal audit appears to have suffered from inadequate staffing, lack of experience, and too little focus on foreign exchange trading as a risk area. Internal audit devotes at most two full-time auditors to auditing all of treasury. Neither of those treasury “specialists” in recent years has had a background or training in trading activities, let alone foreign exchange. The treasury audit responsibilities rest with the same team responsible for trusts (another important audit area), and the manager of that team appears to have had little trading expertise and to have done little to supervise the few treasury auditors he did have. (Indeed, this audit manager appears to have failed even to initial the work papers for the last trading audit.)

Beyond audit, there are other staffing problems. The entire risk assessment department only amounts to two people who are responsible for assessing risk company-wide at Allfirst. And treasury risk control devoted only one full time employee to measuring trading risk in the foreign exchange portfolio. She was extremely inexperienced and appears to have received little support or supervision from others in treasury risk control”. (Promontory, 2002, p.18)

Figure 11 could be extended to include the mass of other similar information that is available to investigators. This would, however, reduce the tractability of diagrams that are already complex. Again the decision about the level of detail to introduce into these figures must be the result of negotiation within the investigatory team. Equally, there must also be some clear mapping between the nodes in the V^2 diagram and the supporting evidence. In previous work we have done this by including unique reference numbers with each vulnerability or violation that can then be cross-references to individual documents gathered as evidence (Johnson, 2003).
Figure 11: V² Diagram Showing Problems in Responding to Reports of Control and Risk Issues
Figure 11 also shows further consequences of the resource constraints imposed on the Allfirst internal audit. The division of one audit team between trust and trading together with problems in the management of these diverse activities led to inadequate oversight for the audit process. At the same time as senior management were becoming aware of the currency feed problems, Rusnak was also exceeding his credit line limit between AIB and UBS. These audit problems partly explain the failure of middle and back-office staff to follow up the reasons for Rusnak exceeding the credit limits. The middle office and credit groups were unsure about who should investigate these problems and in this confusion more credit violations continued to ‘pile up’. The lack of thorough audit and the failure to follow-up on these violations partly explains why they continued to be ‘diagnosed as trader error’ rather than as symptoms of a security violation.

Figure 12 goes on to show the events and contributory factors that led to the discovery of Rusnak’s activities. It is important to study this process of discovery. Previous sections have argued that we are unlikely ever to be able to eliminate potential vulnerabilities in security-critical systems. It, therefore, follows that we must learn as much as possible about those defenses that eventually lead to the detection of particular violations. In this case, there is a link between the $V^2$ diagram and the previous Figure 10 through the continuation symbol labeled 7. The earlier diagram showed that Rusnak had continued to sell year-long ‘deep in the money’ options. These activities trigger a report from a market source to AIB’s Chief Executive that Allfirst is involved in heavy foreign exchange trading. As can be seen at the top of Figure 12, the Allfirst Treasurer responds that there have been no unusual transactions after asking for daily reports on the Allfirst daily foreign exchange transactions. The memo from the AIB Chief Executive was not passed to other senior managers in that bank. After the Treasurer’s response from Allfirst, the matter is dropped.

The $V^2$ diagram in Figure 12 illustrates a further way in which Rusnak’s activities might have been discovered. At the end of the 2000 financial year, Allfirst were required to prepare a variety of financial statements. The Allfirst internal audit group questioned the head of treasury funds management on whether Rusnak’s use of the balance sheet was justified by the profits that he was able to generate. AIB group’s financial reporting unit raised similar questions. As we have seen before, many in the Allfirst senior management were strongly supportive of Rusnak’s trading strategy. The explanation that this was assumed to be low-risk together with the lack of any additional questions from fellow traders and the lack of any systematic review of the previous reports in Figure 11 about poor control strategies all contributed to the internal audit decision to drop their investigations. Similarly, the Allfirst controller, director of finance and head of treasury all meet to allay the concerns raised by the AIB financial reporting unit.

Rusnak’s continued options trading were eventually mentioned in a letter to Allfirst from the Security and Exchange Commission. The Allfirst financial reporting unit found that the large offsetting positions created by Rusnak were a potential source of risk. At the same time, AIB requested a report on Allfirst’s activities for the Central Bank of Ireland. AIB then learn of the increasing foreign exchange transactions and call the Allfirst treasurer. The treasurer then ordered a further investigation. This elicits the response shown as a violation in Figure 12 ‘Rusnak argues the reports are incorrect using trade dates and not year end values’. Again this line of investigation seems to falter. However, together with the lines of enquiry mentioned above, it does form part of a growing suspicion about the trader’s activities.

The final detection factor in this $V^2$ diagram is prompted by the discovery of unconfirmed exchange tickets by back-office staff. Normally exchange options are marked on tickets that are then annotated to indicate that they have been successfully confirmed as ‘legitimate’ with the named counterparties. The supervisor who noticed these tickets then asked their staff to gain confirmation, which had not been usual practice for Rusnak’s trades as explained in Figure 8. The supervisor is eventually told that the trades with Asian counterparties are bogus. Meanwhile as a result of the Allfirst Treasurer’s previous request for daily reports on exchange transactions, he notices a spike in exchange trading that can be linked to Rusnak’s activities. He, therefore, proposed to Rusnak’s supervisor that his positions be closed. These two lines of investigation combine in the continuation symbol 8 that provides a link with the subsequent $V^2$ diagram in Figure 13.
Figure 12: A V² Diagram of the Process of Discovery
Figure 13: A V^2 Diagram of the Rusnak ‘Endgame’
Figure 12 illustrates the start of the discovery process. One of the back-office supervisors finds Rusnack’s unconfirmed option tickets and discovers that they denote bogus trades. At the same time, the Allfirst treasurer becomes aware of spikes in the bank’s foreign exchange trading that he thought had been brought within tight limits. Figure 13 continues the analysis. A can be seen, the supervisor’s senior manager called Rusnack to notify him that they cannot confirm the trades with the counterparties. Rusnack delays the investigation by a violation labeled ‘Rusnack says he will call the brokers to obtain confirmations over night’. At the same time, he created a folder on his personal computer entitled ‘fake docs’. This was subsequently found to contain counterfeit logos and other information relating to the supposed counterparties for the various option transactions.

The $V^2$ diagram goes on to show the events that led from Rusnak’s delivery of twelve apparently ‘confirmed’ option slips to the back-office. The back-office manager believed them to have been forged and so decided to consult with both Rusnak and his superior. The back-office manager argued that the trades should be confirmed by telephone at which point Rusnack became angry and threatened to quit. It is important to reiterate that these events are just as relevant to an investigation into a security violation as the technical and managerial vulnerabilities that created the opportunity for the fraud. As we have seen, previous warnings had been overlooked or ignored. Even at this relatively late stage, it might have been possible for many aspects of the fraud to go undetected. For instance, Figure 13 denotes that Rusnak’s supervisor was concerned that he would quit if he were pressurized too much about his options trading. These concerns partly stemmed from the fact that back-office jobs would be threatened if his trader resigned. These concerns represent a potential vulnerability that could have persuaded the middle management to ignore the warnings they had received about Rusnak’s activities. Rusnak’s supervisor also argued that confirming trades was a back-office problem. Again, this response may have been motivated by the estimated $300,000-$500,000 that it would cost to close his positions. It may also have been motivated by the personal support that the manager had provided for his traders supposed activities in previous years. Rusnak’s supervisor agrees that the confirmations looked bogus but asked the back-office staff to again seek confirmation over the phone.

Rusnak later returned to the meeting between the back-office manager and his supervisor. He offered help to confirm the transactions. However, it is Friday and the Asian markets will be closed until Sunday midday. Rusnak promises to give them the broker’s telephone numbers by 21:00. The call is never made. A back-office employee rings Rusnak on Sunday afternoon asking for the confirmations and their associated telephone numbers but cannot reach Rusnak. Rusnak does not appear at his desk the following Monday. His supervisor and the senior back-office manager then report the bogus transactions to the Allfirst treasurer. The treasurer joined Rusnak’s supervisor in driving to the trader’s house but they find that he has left. The Allfirst treasurer then passes his concerns on to others in the senior management of Allfirst and of the AIB group.

The previous pages have shown the way in which $V^2$ diagrams can be used to map out the events and contributory factors, the violations and the vulnerabilities that characterize serious security incidents. The intention has been to provide a detailed case study so that this approach might be extended to other adverse events. This approach also helps investigators to focus on the detection factors that combine to help organizations identify that they may have a potential problem. In Rusnak’s fraud there were several opportunities where his violations might have been exposed. These range from external reports, such as market sources questioning the extent of foreign exchange dealing at Allfirst through to regulatory intervention, such as the questions asked in response to the report required by the Irish Central Bank. Staff vigilance also played a role. Even though the Allfirst internal audit teams were ill-prepared to identify Rusnak’s actions they did notice problems in the currency feed. As we have seen, however, the $V^2$ diagrams map out the various factors that combined to divert or extinguish these lines of enquiry. Key personnel had significant investments, in terms of time and reputation, in the success of Rusnak’s activities. They were also aware that the future of their own careers and those of their colleagues depended to some extent on the trader’s operations. At other times, several members of staff decided to take personal responsibility for investigating their concerns rather than asking more senior management to conduct a more sustained enquiry. Finally and above all, the links between audit and risk management were never clearly established. Doubts about the accuracy of the key VaR metric and about the security of the currency feeds never triggered the sustained audit that might have disclosed the fraud at a relatively early stage.
Figure 14: A V² Diagram of Software Issues
The previous $V^2$ diagrams have focused on the construction of an event-based model of the Rusnak fraud. There are other ways in which this technique can be used. Diagrams can also focus in on particular aspects of a security-related incident. For example, Figure 14 shows how a $V^2$ diagram can be constructed to look more narrowly at the role that software-based systems played in the fraud. This is particularly important given continuing concerns about the management and oversight of access provided by this class of applications. The extension symbol labeled 2a refers back to Figure 6. This described some of the contextual factors that stemmed from the merger between Allfirst and AIB. In particular, it relates to AIB’s decision that Allfirst should be allowed considerable independence and that the new acquisition should be managed with a ‘light hand’. AIB had been one of the first banks to invest in a software system called Opics. The Opics application automates and centralizes a number of back-office functions. It can also be used in conjunction with a ‘sister-application’ known as Tropics that supports currency trading. An important benefit of using these applications together is that they can enforce a separation of back-office and front-office activities. They can also be used to trace the confirmation of options that were created by the front-office staff and should have been monitored by back-office employees. Tropics was not installed at Allfirst. Hence the software did not support the tracking and clear division of responsibilities that might have prevented many of the vulnerabilities and violations that were identified in previous $V^2$ diagrams.

As can be seen in Figure 14, the decision not to install Tropics was justified on many grounds. Firstly, the costs of the software may not have been justified by the relatively small size of the trading desk. Also, at the time of merger AIB appeared to be happy with the Allfirst risk control and trading statements. They arguably did not see any justification for the additional monitoring facilities provided by the Tropics application. The decision to invest in Tropics can also be partly explained by a failure to learn from the Barings experience where a trader had managed to erode the separation between front and back office functions. Finally, there was no tradition for preserving this separation in terms of the electronic systems that support the work of Allfirst staff. The outcomes from the decision not to install Tropics included the lack of any automatic confirmation for trades. The decision not to install Tropics also prevented any automatic warnings for traders when their activities exceeded credit limits.

In addition to the Opics and Tropics systems, Allfirst might have been protected by the introduction of the Crossmar software that was used by AIB. This application also provided automated confirmation for trades using a matching service. Allfirst did not use the Crossmar software and so most of the confirmation relied upon back-office staff to fax requests to overseas markets. This manual confirmation was vulnerable to interruption and dislocation due to overseas trading hours. It was also open to pressure from traders such as Rusnak. Although we have not included it in the current analysis, Figure 14 might also be extended to illustrate the additional pressures that Rusnak’s activities created for the back-office staff. His bogus options relied upon the continual generation of additional transactions beyond his legitimate trading activity. One side-effect of the fraud would, therefore, have been to increase the workload on back-office staff which in turn may have left them even more vulnerable to attempts to delay or ignore confirmations on a rising number of trades. AIB had also decided to exploit a software application known as RiskBook. This uses front and back-office systems to calculate the bank’s risk exposure. Previous sections have described how Rusnak was able to affect the VaR calculations and there is reason to suppose that the use RiskBook might have offered some protection against these actions. Allfirst were not, however, part of the first roll-out for the RiskBook software within Allfirst. It is deeply ironic that Rusnak had been asked to specify the requirements for this new risk management software.

CONCLUSIONS AND FURTHER WORK

A number of commercial and governmental organizations have recently argued that we must look beyond the immediate events that surround security-related incidents if we are to address underlying vulnerabilities (Austin and Darby, 2003). It is important to look beyond the immediate acts of ‘rogue traders’ or individual employees if we are to correct the technical and managerial flaws that provide the opportunities for security to be compromised. This paper has, therefore, provided an introduction to Violation and Vulnerability analysis using $V^2$ diagrams. The key components of this technique are deliberately very simple; the intention is to minimize the time taken to learn how to read and construct these figures. The paper has, in contrast, been motivated by a complex case study. The intention has been to provide a sustained example at a level of detail that is appropriate to an initial investigation into complex security incidents. Previous pages have provided a sustained analysis of Rusnak’s fraudulent transactions involving...
the Allfirst bank. This case study is appropriate because it involved many different violations and vulnerabilities. These included failures in the underlying audit and control mechanisms. They included individual violations, including the generation of bogus options. There were also tertiary failures in terms of the investigatory processes that might have uncovered the fraud long before bank personnel eventually detected it.

Much remains to be done. We are currently working with a number of organizations to extend and tailor the techniques in this paper to support security investigations in a range of different fields, including both financial and military systems. There is a common concern that the $V^2$ approach will provide a standard means of representing and modeling the outputs of an investigation into the causes of security-related incidents. In each case, however, we are being encouraged to extend the range of symbols represented in the diagrams. For example, these might be used to distinguish between different types of barriers that should have led to the identification of a violation or vulnerability. In terms of the Allfirst case study, the decision not to tell senior management about concerns over the Reuter’s currency feed via Rusnak’s PC would have to be represented using a different type of symbol. The intention is that analysts would then be encouraged to probe more deeply into the reasons why this potential warning was not acted upon. An important concern in this continuing work is, however, that the additional notational elements will increase the complexity of what is a deliberately simple approach. It is critical to avoid additional complexity in the analysis of what are almost always extremely complex events.

Further work also intends to explore the use of $V^2$ diagrams as a communication tool with wider applications. In particular, the outcomes of many security investigations must be communicated to diverse groups of stakeholders. These are not simply confined to security professionals and senior management in the target applications. In particular, it is often necessary to communicate findings about the course of an incident with members of the public who may potentially be called upon to act as jurors in subsequent litigation. The complexity of many recent security related incidents makes it vitally important that we find the means to help people understand the events and contributory factors that form the context for many adverse events. Similarly, political intervention is often triggered by incidents such as the Allfirst fraud. It can be difficult to draft effective legislation when key figures lack the necessary time and briefing material to fully follow the events that they seek to prevent.

REFERENCES


Naval Surface Warfare Centre, Dahlgren, Computer Security Incident Handling Guidelines, Department of the Navy, Commanding Officer, Fleet Information Warfare Center, Virginia, USA, 2002.


