# Limits to Effectiveness in Computer Security Incident Response Teams

*By*
***Johannes Wiik  Jose J. Gonzalez***
*Faculty of Engineering and Science*
*Security and Quality in Organizations*
*Agder University College*
*Grooseveien 36*
*NO-4876 Grimstad, Norway*
*Phone: +47 3725 3000*
*Email: Johannes.Wiik@hia.no  Jose.J.Gonzalez@hia.no*

***Klaus-Peter Kossakowski***
*Software Engineering Institute – Europe*[1]
*An der Welle 4*
*DE- 60322 Frankfurt, Germany*
*Email: kpk@sei.cmu.edu*

## Abstract

*In a continuously changing environment, a Computer Security Incident Response Team (CSIRT) has to evolve to sustain or improve its effectiveness. The main task of a CSIRT is to mitigate the effects of computer security incidents. A frequently identified problem is that CSIRTs are over-worked, under-staffed and under-funded. We present a System Dynamics simulation model of such conditions based on a case study. The model is a first attempt to understand the main factors influencing a CSIRT's effectiveness, and to improve its performance. Based on theory from process improvement and information from the case study, we identified that short-term pressure from a growing incident work load prevents attempts for developing more response capability long-term, leading the CSIRT into a "capability trap". Fundamental solutions will typically involve a worse-before-better trade-off for management.*

[1] Dr. Klaus-Peter Kossakowski is Visiting scientist at SEI-Europe

# 1 Introduction

The primary mission of a Computer Security Incident Response Team (CSIRT[2]) is to help other organisations handle incidents occurring in computer networks. CSIRTs may also provide a wider set of services. West-Brown et al. (2003 p. 177) describe (some) main challenges for CSIRTs:

> *"To ensure successful operation, a CSIRT must have the ability to adapt to changing needs of the environment and exhibit the flexibility to deal with the unexpected. In addition, a CSIRT must simultaneously address funding issues and organizational changes that can affect its ability to either adapt to the needs or provide the service itself."*

Current research suggests that most CSIRTs suffer from overstretched resources and a growing work load, indicating that making such organisations effective is a significant challenge. The area of CSIRTs is still very much a pioneering field (West-Brown et al. 2003 p. 179), and there is little research addressing the managerial aspects of CSIRTs. In particular, we have not been able to identify research targeting the causes of the resource and work load problem listed above, nor how to make such organisations more effective.

The results presented in this paper represent a preliminary attempt to gain a better understanding of how a CSIRT can handle a growing work load with limited resources, factors that restrict its effectiveness and ability to improve, as well as a proposed solution to improve its performance long-term. Central to our approach is a System Dynamics model of a case study.

# 2 Survivability and the CSIRT Mission

Despite the fact that CSIRTs have developed over almost two decades, there is still no widely accepted way to classify an organisation as a CSIRT. We will use the definition in the CSIRT handbook: "For a team to be considered a CSIRT, it must provide one or more of the incident handling services: incident analysis, incident response on site, incident response support, or incident response coordination." (West Brown et al. 2003 p. 23) Given this definition, a CSIRT should therefore mainly be considered reactive in nature.

A simple way to describe a CSIRT's mission is: "to minimize the impact of an incident to a company and allow it to get back to work as quickly as possible" van Wyk (2001), or "to be a focal point for preventing, receiving and responding to computer security incidents" (Killcrece et al. 2003a, p. xi). It is the responsibility if CSIRT managers to achieve such goals. There are many options and a wide range of services can be offered by a CSIRT to accomplish their goals. Some of services target (proactive) prevention of incidents, while others minimise the negative consequences of incidents in a more reactive manner.

[2] There are many names used for CSIRTs, such as Incident Response Team (IRT), Computer Emergency Response Team (CERT), etc. For a comprehensive list of alternative abbreviations, please see appendix B of the Handbook of Computer Security Incident Response Teams (CSIRTs), West-Brown et al. (2003).

The first CSIRTs were established to help organisations that had become victims of malicious attacks[3]. Hence, the pioneering CSIRTs offered mostly reactive services, and this is what most CSIRTs still do.

There has been a growing realisation that more proactive services are needed (Killcrece et al. 2003a). In recent years some CSIRTs have included proactive services. Nevertheless, reactive services will always be required, as indicated by the CSIRT definition above.

The new emerging survivability paradigm provides one of the most important reasons for a continued need for reactive services. This paradigm is replacing traditional security thinking, which focuses on achieving full protection against malice from a fortress perspective (Blakley 1997). One of the basic assumptions within the survivability paradigm is that no matter how much security you put into a system, it will never be totally secure (Lipson and Fisher 1999 p. 3). This is why a part of a CSIRT's mission is to mitigate the impact of occurring incidents.

Whatever services a CSIRT provides, management tends to encounter similar problems. Next, we will take a closer look at some of these problems.

# 3 Common Problems Among CSIRTs

Frequently referenced problems in the CSIRT community are over-stretched resources and a need for continuous improvements. Killcrece et al. (2003a p. 128) refer to many such problems, for example: lack of funding, lack of management support, lack of trained incident handling staff, lack of clearly defined mission and authority, and lack of coordination mechanisms. Staff burnout is a common side-effect of over-stretched resources (Killcrece et al., 2003a p. 78). Already in 1994 it was stated: "About the only common attributes between existing Incident Response Teams are that they are under-funded, under-staffed, and over-worked." (Smith 1994)

If we change the perspective to the attacker's side of the problem, we find exogenous factors that may influence the daily operation of a CSIRT. According to Lipson (2002, p. 9), "although the sophistication of Internet attacks has increased over time, the technical knowledge of the average attacker is declining, in the same manner that the technical knowledge of the average user has declined". Hence, not only are more people able to launch attacks, the scope and frequency of such attacks is growing. The volume of attacks is thereby continuously increasing (Killcrece et al., 2003a p. 113).

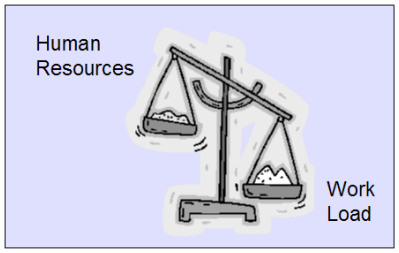


**Figure 1: A fundamental problem for a CSIRT is to balance a growing work load with limited human resources.**

This would indicate that the need for CSIRT services is growing, but, as seen above, many CSIRTs are drowning in an increasing work load and will continue to be overwhelmed unless they are able to adapt to the changing environment.

The fundamental problem can be portrayed as in Figure 1: How can a CSIRT handle a growing work load and stay effective with limited resources?

[3] The first CSIRT was CERT Coordination Center (CERT/CC) at Carnegie Mellon University as a response to an Internet worm incident in 1988. CERT/CC has later developed into a generic centre of Internet security expertise. For more information, s. www.cert.org.

# 4 A System Dynamics Approach[4]

None of the problems listed above should be studied in isolation. A CSIRT interacts with its surrounding environment that includes, for example, its constituency, funding institution, other CSIRTs, etc. Also, there are many internal interactions between different services, and some services may even share the same resources. Typically, the problems seen today have accumulated over time. Consequently, we are dealing with a complex and dynamic problem.

The System Dynamics method is a promising approach for addressing such problems. System Dynamics is based on the theory of nonlinear dynamics and feedback control that come from mathematics, physics and engineering (Sterman 2000). Tools are available to make quantified models that can be simulated. As such tools are applied to social systems with human decision making and human behaviour, System Dynamics also incorporates theories from cognitive and social psychology, economics and other social sciences (Sterman 2000). Consequently, a System Dynamics approach is inherently multi-disciplinary.

In many cases, researchers study the trend of events over time to understand how a system will evolve. System Dynamics emphasises that a deeper understanding of system behaviour is possible. A basic assumption in System Dynamics is that the initial state and the causal feedback structure of a system determine its behaviour over time (Sterman 2000). As the system evolves over time, the behaviour of the system might alter the dominant structure. Put in a different way, the state of a CSIRT influences the decisions we make. The decisions we make will alter the state of a CSIRT, and hence, feed back and influence our future decisions. Such a system perspective means that System Dynamics takes a broad and holistic perspective where all of the main factors are understood, and the time frame under consideration is long enough to understand how problems have emerged and how they can be solved.

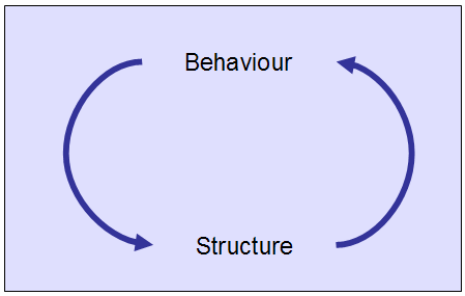


**Figure 2: System Dynamics assumes that the structure of a system determines its behaviour. As a system evolves, the dominant structure may also change.**

When we try to elicit information about the structure from experts, such information is never more than a set of assumptions. Often, even the very best and knowledgeable experts will be reluctant to share their understanding of a system. Not because they do not want to, but because they are uncertain about the main factors in an ambiguous world. The only way to test the assumptions of such a model is through simulation (Sterman 2000 p. 27). The ability to simulate a model makes it possible to test its structural assumptions. Simulation also allows one to artificially speed up time, and years can be simulated in a matter of seconds. Hence, delayed feedback effects are much better understood in a controlled model environment compared to the real world. Often, simulation reveals that the initial assumptions were wrong, because they do not mimic the reality we know. Hence, we reveal that our initial understanding of the system was wrong, which again will lead us to find alternative explanations for past performance. For evaluating future scenarios, simulation also allows for experimentation with policies that would otherwise be impossible both for practical and ethical reasons. Hence, we can test policies under different conditions and create a holistic

[4] Sections 4-6 contain elementary issues in System Dynamics for the benefit of the information security community.

understanding of possible scenarios for the future. A simulation model can never predict the future, but it can significantly enhance our understanding of how the future might look like.

# 5 Information and System Dynamics Modelling

System Dynamics modelling is dependent on information. However, it is very often difficult to gain access to security information. There are many reasons for this. Firstly, it can be due to strict information policies and the sensitivity of data. The disclosure of information regarding incidents is seen as potentially damaging to customer reputation or even to the reputation of the CSIRT itself (West-Brown et al., 2003). Secondly, it can be difficult to get information about hackers, as they are experts in hiding their tracks. Thirdly, after some time, information stored may not be considered relevant for the CSIRT. Due to the sensitivity concerns mentioned above, information is typically disposed of in such a secure manner that it cannot be retrieved again. And fourthly, information that can be useful in a study such as ours is simply not stored at all as it may not be perceived as useful for the CSIRT's daily operations.

Consequently, nearly all of the information available about CSIRTs today is anecdotal. To the best of our knowleddge, the most extensive information publicly available, with respect to how a CSIRT operates, is "State of the Practice of CSIRTs" (Killcrece et al. 2003a). Still, as stated by the authors of this report, the information gathered from surveys among CSIRTs for the report is still quite limited and should be treated as such (Killcrece et al. 2003a p. 7).

According to Forrester (1994) there are three sources of information that are useful in modelling:

- the mental database,
- the written database, and
- the numerical database

Indeed, the most useful information for modelling is not necessarily restricted to numerical data. "Effective model building must draw upon the mental database" (Forrester, 1994 p. 73) because it contains knowledge about policies and structure. To understand system behaviour, one must understand the underlying structure, and much of the information about the structure comes from the mental database. The written database, such as articles, books and other publications, provides useful information, but the problem is that we cannot query a book, only the author. Hence, written information is less rich than the mental database. Even more limited is the numerical database, although such information can be useful to quantify parameters in a simulation model. Also, one of the most common uses for numerical data is to compare time-series data from real world history as a part of the validation process of a simulation model. In our research, we have used all three types of information – mental, written and numerical.

Due to the limited availability of numerical and written data, we opted for a more inductive approach where we employ a case-study to elicit as much information as possible from the experts in the field.

Most of the information has been gathered through interviews and in meetings. The questions were presented in advance, and the answers were then put down on paper and later reviewed by management to check that the information was correct, or to add more information not discussed in the interview. The meetings typically took place at a later stage when we had a running model. Consequently, we were able to simulate and quickly validate many of the new

assumptions we identified leading to new meetings and discussions about assumptions. We have also gained access to some time-series of numerical data that have proven to be very useful in the validation of the model.

Through cooperation between Agder University College[5] in Norway and DFN-CERT[6] in Germany, a PhD project has been initiated. The University of Bergen[7] in Norway is also involved in this PhD project. The purpose is a better understanding of the factors limiting the effectiveness of a CSIRT to identify robust policies that can sustain or improve their effectiveness[8]. DFN-CERT was originally established as a research project within the German Research Network (DFN) in 1993, and in 1999 it was transformed into an independent company still serving the same constituency.

DFN-CERT is an external, centralised and coordinating CSIRT according to the definitions in the handbook Organizational Model for Computer Security Incident Response Teams (Killcrece et al. 2003b). It is external to the organisations in the constituency, and works as a centralised point of contact for incident response in the German research network. However, it has no authority over its constituency, which thereby restricts its role to mostly that of giving advice to victims of incidents rather than actually doing the recovery work or enforcing specific security measures. In addition, DFN-CERT performs a coordinating role by addressing sources of attacks. This can be done, for example, by contacting other CSIRTs if an attack was launched from their constituency. Hence, the insights presented here are not necessarily easily transferable to other types of CSIRTs. However, we do expect that other coordinating CSIRTs face similar problems, and consequently, some of the insights may be valuable to such organisations. DFN-CERT is one of the oldest CSIRTs in operation. It is, therefore, a good case for studying the effectiveness of a CSIRT from the perspective of different managers and strategies. Thereby we can compare the dynamics of different strategies under different management regimes.

Next, we will present the simulation model developed from the case study of DFN-CERT.

# 6 Presentation of the model

## *6.1 The Modelling Process*

A System Dynamics approach typically involves an iterative process following several steps[9]:

1. Problem formulation
2. Model formulation
3. Model validation and testing
4. Policy experimentation and future scenarios

[5] Agder University College is the funding institution for the PhD project.
[6] DFN-CERT is an abbreviation for Deutsches Forschungsnetz – Computer Emergency Response Team. Its constituency is the German Research Network (DFN). For more information see: http://www.dfn-cert.de/
[7] The University of Bergen is formally responsible for the PhD education.
[8] The PhD project is anchored in the research cell "Security and Quality in Organizations" at Agder University College headed by Professor José Gonzalez. Other papers with a specific CSIRT focus include Gonzalez (2005) and Sawicka, Gonzalez and Qian (2005), Wiik and Kossakowski (2005).
[9] Sterman (2000) gives an in-depth explanation of the modelling process using a 5-step approach. For simplicity we use a 4-step approach here as we mainly try to explain the insights of the model without going into too much detail about the approach.

Throughout this modelling process, previous steps are often redefined as new knowledge about the problem and the system under investigation is identified. In the following sections, we will describe the model and the model insights following these steps.

## 6.2 Problem Formulation

The typical approach for making a System Dynamics simulation model starts with a problem. A well-defined problem helps us to consider what is relevant to include in a model and what is not. In our case, such a problem can be described as the development of some key variables of concern to a CSIRT over time. The simulation model we have developed is quite extensive and actually addresses a wide range of issues faced by a CSIRT such as the effectiveness of both proactive and reactive services. The main problem we address in this paper can be described in the two time-graphs displayed in Figure 3. In this figure, we see the historical development of the human resources allocated to reactive work, and the total number of incidents handled, as well as the split between high and low priority incidents. The work load is also shown with a different scale in Figure 4.

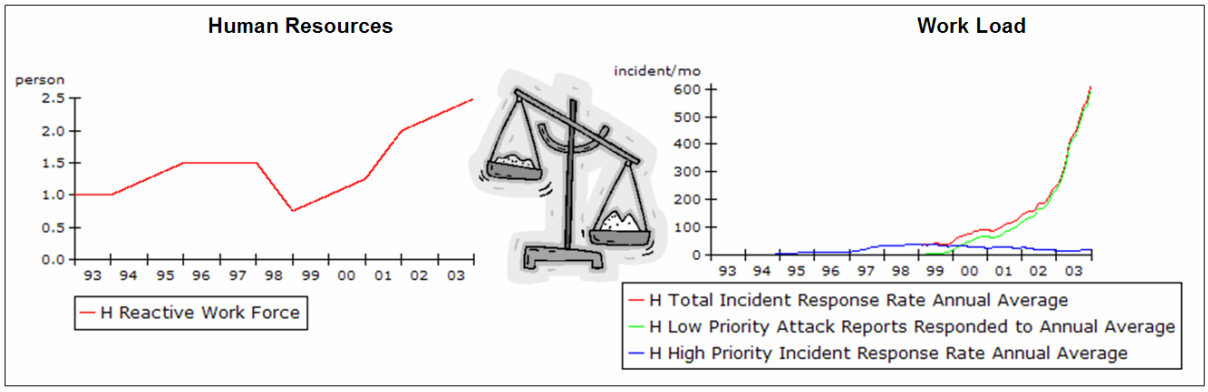


**Figure 3: The historical development of some key variables**

We comment Figure 3: First, the available *Reactive Work Force* has not grown as much as the work load. Second, there is a period within 1998 when available resources were actually decreasing. And third, both graphs reflect a significant shift starting in 1999.

The main reason for the first fact was limited funding for human resources was limited. There were no business models that would allow for the exponential growth we observe on the right graph, particularly as this graph reflects the growth of low priority attacks. We will argue, however, that such attacks need to be considered, and that some of them actually help to identify compromised systems and incidents that need more attention, but most of them are denied anyway and originate outside the constituency.

The second fact reflects a huge turnover in human resources (not unusual, but certainly undesirable). Due to attractive offerings from industry, together with limited abilities to provide higher salaries and permanent employment within the university environment, experienced staff members left their positions, and it was difficult to attract new staff members in time to fill all open positions within the transition period. This was actually the catalyst responsible for the transition from a university project towards an independent legal entity as non-profit company not bound by university policies

The third fact is owing to a change in support policies. Before 1999 low priority attacks, such as port scans or open mail relays causing spam and other similar issues, were not handled routinely; they were only responded to if a greater impact on the constituency was detected.

By changing policies, i.e. by offering support for low priority attacks after 1999, the overall work load was adversely impacted by the exponential growth of these attacks.

Compared to the growing number of attacks, the number of more severe incidents that needed intense support from the team, has been relatively stable (see Figure 4) – although they have changed from a technical point of view. The numerical stability is due to the existing awareness and available security measures, which only leaves a limited number of incidents to be handled manually by the local staff of the organizations. In such cases, the team provides the expertise not available locally and enables the organization to mitigate the actual problem. The issues related to the handling of such severe high priority incidents are also within the scope of this PhD-research project but they are not addressed in this paper.

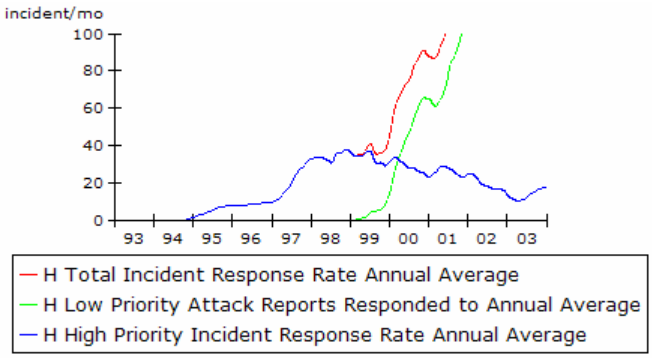


**Figure 4: The rate of high priority incidents handled has been relatively stable over time compared to the low priority growth since 1999.**

The common means for addressing the increasing work load has been to work harder, reduce slack, and work overtime. More fundamental policies, such as attempts to work smarter and enhance the incident response process have been short-lived and ended in failure.

The growing mismatch between resources and work load begs the questions :

- What factors have influenced the CSIRTs effectiveness over time?
- How can we improve the effectiveness of the CSIRT in the future?

Although our simulation model has been built to understand effectiveness of both reactive and proactive services over the entire lifespan of the CSIRT since 1993, we will in this paper, restrict ourselves to:

1. The period after 1999 mainly
2. The growth of low priority attack reports after 1999, and
3. Incident response (we will for the time bing mostly disregard the proactive effort).

Any other aspects will only be included as we deem necessary to clarify the assumptions in the model.

## *6.3 Model Formulation*

Through our case study, we have found an interesting resemblance between our CSIRT-related problem and process improvement topics studied at the Massachusetts Institute of Technology (MIT), together with some collaborating institutions (Keating et al. 1999, Repenning and Sterman 2001, and Repenning and Sterman 2002). Large portions of our preliminary findings have been inspired by their work. We will therefore refer to their work as we describe the simulation model.

For simplicity, we present here a simplified conceptual version of the model. The full simulation model, with around 300 variables, is too complex for the scope of a single paper.

### 6.3.1 Two Accumulation Processes That Must Be Balanced

After first defining the dynamic problem in the case, we need to understand how the underlying structure drives the problem behaviour. The System Dynamics method uses a simple graphical language to describe the structure of feedback and accumulation processes. In Figure 5 below, we present two examples of accumulation processes using System Dynamics' stock and flow notation.

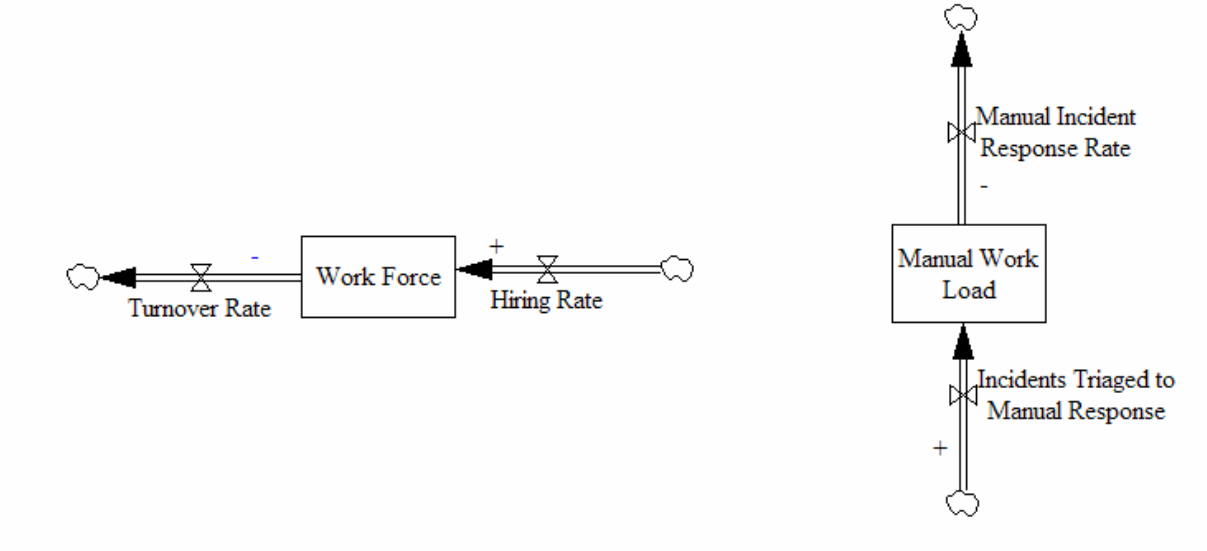


**Figure 5: Describing accumulation processes using stocks and flows**

To the left we assume that the *Work Force* is a stock. Stocks are accumulations (Sterman 2000). A stock can never change instantly – only through its associated flows. As stocks do not change instantly, they represent delays in a system that separate cause and effect, and allow for disequilibrium in systems. For example, it takes time to hire people and add them to the work force. The level in a stock represents the state in a system at any given point in time, while the flows represent the changes to the states. Hence, the inflow to the *Work Force* is the *Hiring Rate*, while the outflow is the *Turnover Rate*.

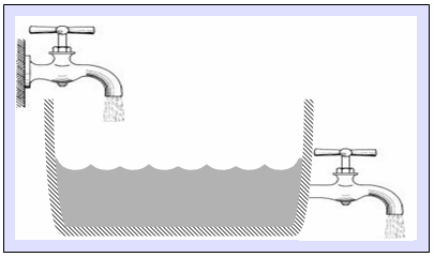

**Figure 6: An easy way to understand the concept of stocks and flow is to use the example of a bathtub. The level of water in a bathtub (stock) is regulated through the rate of water through the faucet (inflow) and the rate through the drain (outflow).**

A typical example that illustrates the concept of stocks and flows is how the level of water in a bathtub (a stock), is regulated through its inflow and outflow of water through the faucet and the drain as illustrated in Figure 6.

To the right in Figure 5, we see a different example where information about attacks is reported through the inflow (*Incidents Triaged to Manual Response)* and accumulates into a stock (*Manual Work Load)*. Note that the Manual Work Load is not the same as the number of open incidents, but rather the information gathered and considered for response for the time being. As this information is processed, represented by the outflow (*Manual Incident Response Rate)*, the *Manual Work Load* will decrease.

### 6.3.2 Adding Feedback to the Stock and Flow Structure

So far we have only described some of the stocks and flows in the model. Next, we will take a closer look at another aspect of the causal structure, namely the information feedback-loops connecting the stocks and flows. Consider Figure 7 where we have added some information links and some new auxiliary variables[10]:

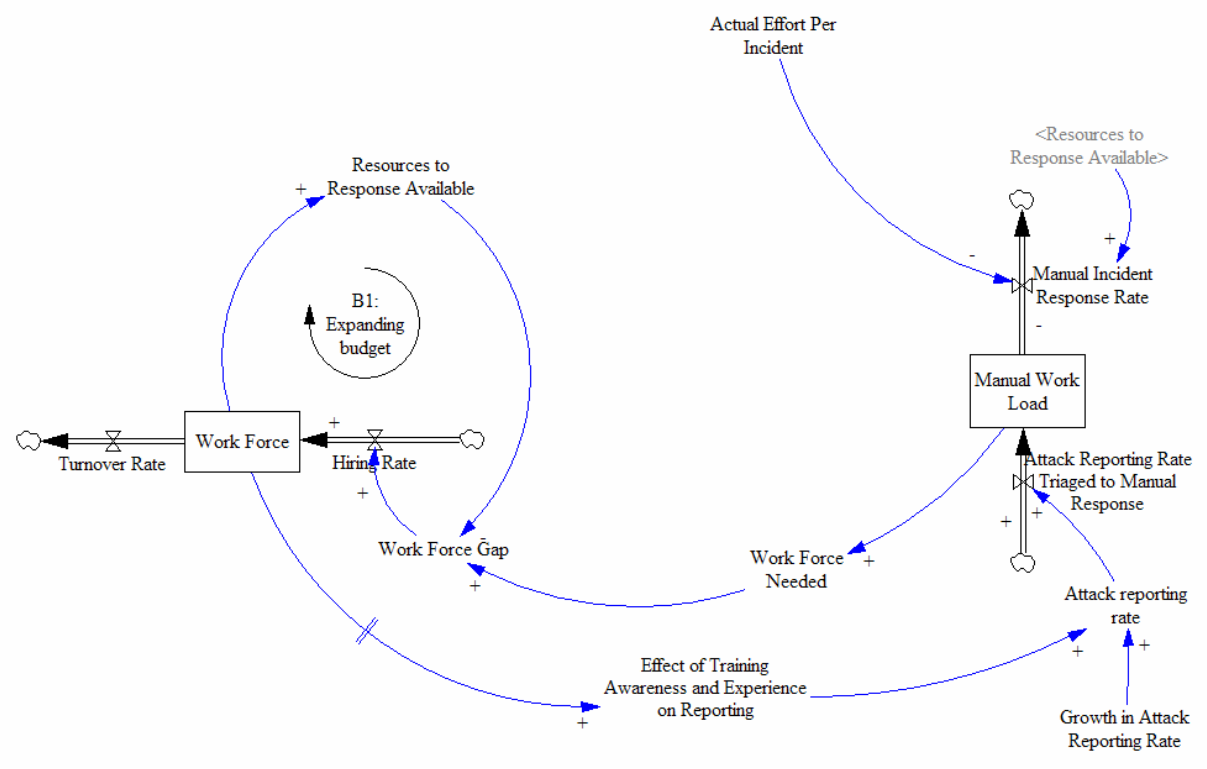


**Figure 7: Introducing feedback in the model**

The size of *Manual Work Load* determines *Work Force Needed*. If *Resources to Response Available* increases compared to *Work Force needed*, *Work Force Gap* will increase too. Note the +/- signs for the information links. If a cause variable changes (increase or decrease) a positive link means that the effect changes in the same direction, all else equal. Conversely, a negative link means that the effect changes in the opposite direction, all else equal. Also note that the variable *Resources to Response Available* is used twice in the diagram to avoid too many crossing information links. One is the original, while the other is a snapshot (meaning the same variable) of the original.

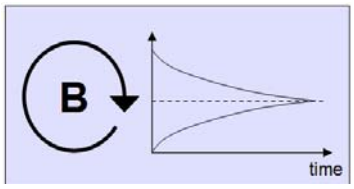


**Figure 8: A dominating balancing feedback loop will generate converging behaviour towards a goal. In this example the goal is constant (dotted line). A goal may also be dynamic.**

Management can close the *Work Force Gap* through the *Hiring Rate*. Thus we have identified a balancing feedback loop that tries to reach equilibrium where the workforce and the work load are balanced to each other. However, funding is limited, and thereby hiring new people is difficult for most CSIRTs, so this loop will normally have a limited effect. It is, however, the most important loop described in Figure 7.

[10] Strictly speaking, the mathematical description of a System Dynamics model only requires stock and flow variables (Sterman 2000). Auxiliary variables are added to clarify the intermediate calculations between stocks and flows, and hence, enable communication of the model structure to a wider audience.

Another important causal relationship in Figure 7 is the link from *Work Force* to an aggregated variable labelled *Effect of Training Awareness and Experience on Reporting*. This variable represents the effect of using resources on awareness building and training in the constituency. In addition, the accumulated experience will influence the reputation of the CSIRT, as people with experience have built up relationships with other communities, etc. All these factors influence *Attack Reporting Rate*, both from the constituency and from external parties. Another new symbol in this section of the model is the delay-mark, which is illustrated with two short lines crossing an information link. Such a symbol indicates that there are the causal effect propagates between the two variables with a significant time delay.

Also note the variable *Growth in Attack Reporting Rate*. This growth factor has been significant in DFN-CERT since 1999 when the definition of an incident was expanded to include attacks such as port-scans and spam. We have assumed that the main reason for the growth is exogenous to the CSIRT, and accordingly, we have modelled this as an external factor to the model. Similar exponential growth patterns can be seen in the aggregated CERT/CC statistics[11] as well, and the main attribution for this is the growing automation of attacks on the internet over time. In addition, most customers will have firewalls installed, and the automatically generated firewall logs will be well-suited to discover such attacks and they can easily be forwarded as information to the CSIRT. We mention also that as the CSIRT has committed itself to participate in cooperation with other organisations for exchanging information it has also linked itself into infrastructures for reporting. These infrastructures may not lead to a growth in themselves, but they will facilitate such growth by making reporting easier.

As mentioned in the dynamic problem definition above, the exponentially growing work load, and the corresponding resources required to handle this work load, have not followed the same pattern over time. In other words, the balancing loop, “B1: Expanding Budget”, has not been the main balancing factor. Next, we will investigate what factors might instead have played a main role in balancing the growing work load.

### 6.3.3 Balancing Feedback Compensates for the Growing Work Load

Consider how the growing incident work load has been handled in the past: Incidents are reported to the CSIRT by its constituency. With a growing work load, the workforce will find other ways of handling more incidents. A typical quality measure for a CSIRT is its response time. Most CSIRTs try to respond within a certain time-frame to reported incidents (West-Brown et al., 2003 p. 65). The accumulated number of incidents reported, the desired response time, and the available resources for response will thereby determine *Effort per Incident Needed*.

This means that staff working with incident response has to be more efficient as *Attack reporting rate* increases. There are several ways to accomplish this, for example:

- The workforce can try to work overtime and reduce slack time to handle more incidents during a single working day.
- Simple scripts can easily help filter out the appropriate information.
- Similar incidents can be clustered together to handle several at the same time in the same manner.

[11] http://www.cert.org/stats/#incidents

- The time used to follow up a customer can be reduced. Instead of calling up a customer that does not respond, an easy solution is to close the incident if no one on the customer side responds to e-mails from the CSIRT.
- By updating the contact information, the CSIRT will spend less time and effort identifying the right point of contact during incident response.

These are just a few examples of how incident response personnel can become more efficient. In practise, the CSIRT will gradually adjust *Actual Effort per Incident* to match *Effort per Incident Needed*.

This structure constitutes a new balancing loop, "B2: Working Harder", and it is one of the main factors that enables the CSIRT to balance the growing work load as well as the incident handling effort (See Figure 9).

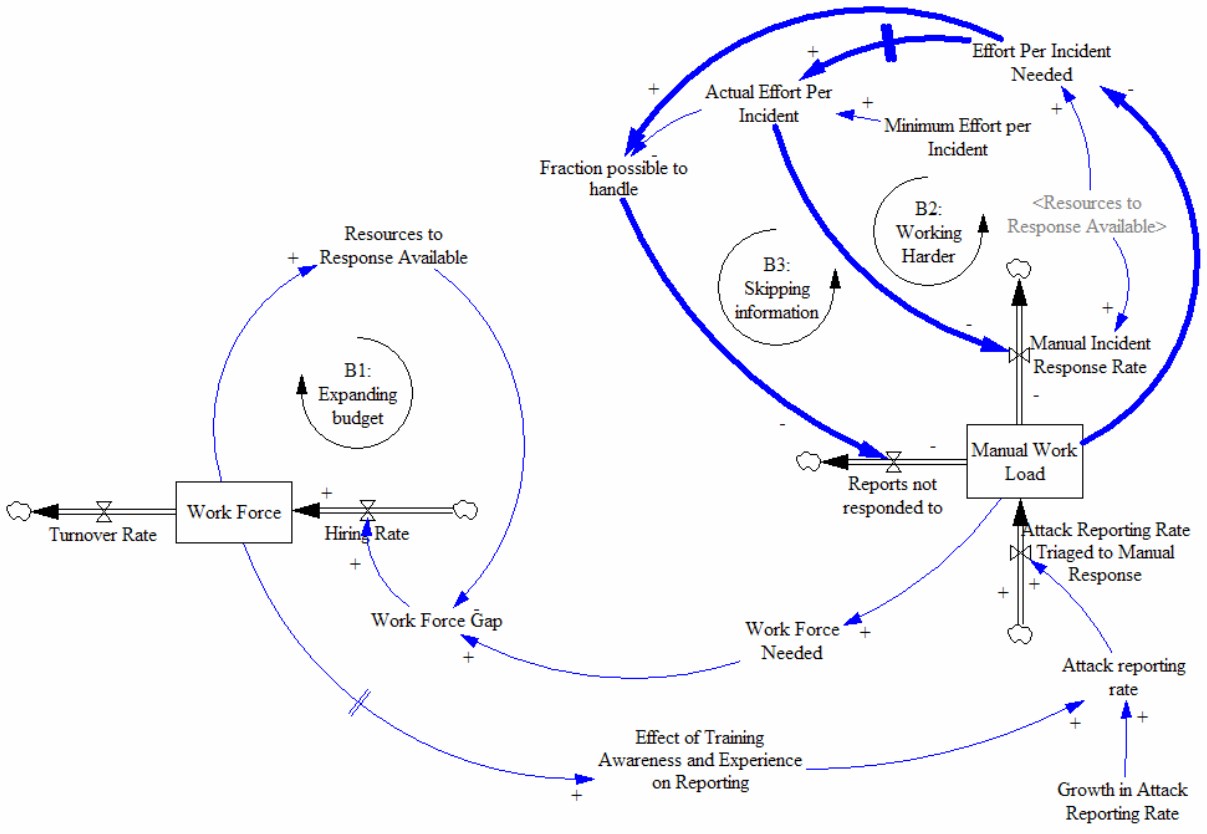


**Figure 9: Compensation for the exponential growth in attack reports**

However, there is a limit to how much a person can reduce *Actual Effort per Incident* by working harder. This limiting factor is captured in the variable *Minimum Effort per Incident*. As the limit is approached and surpassed, a growing gap between *Effort per Incident Needed* and *Actual Effort per Incidents* starts to emerge. Hence, as the *Fraction Possible to Handle* decreases the staff has no other option than skip information that has been reported (they do not follow up all reports about low priority attacks). A possible way to prioritise under such conditions is, for example, to accumulate similar types of reports referring to the same incident, and as the accumulated reports exceed a certain threshold the priority is increased, and the attack reports are responded to. Less frequent reports about attacks are not responded to or used in the response process, modelled as the outflow variable from the Manual Work Load, *Reports not Responded To*. This outflow closes a third balancing loop of importance, "B3: Skipping Information".

Gradually a shift from B2 to B3 will take place when the CSIRT is no longer able to balance the work load. Naturally, managers of a CSIRT will start looking for solutions to this problem, and hopefully long before the limits of B2 are reached.

### 6.3.4 Working Smarter: Investing in Automation

A CSIRT can also work smarter. Parts, but not necessarily all, of the incident response processes can be automated by developing tools that take reports as inputs. An automatic response will then follow, providing appropriate self-help information to the victim. In addition, other tools can help filter the incoming information more efficiently, and to identify incidents in a more efficient manner and on a more aggregated level. This is typically the case with lower priority attacks, but it is less applicable to more complex and higher priority incidents such as root compromises.

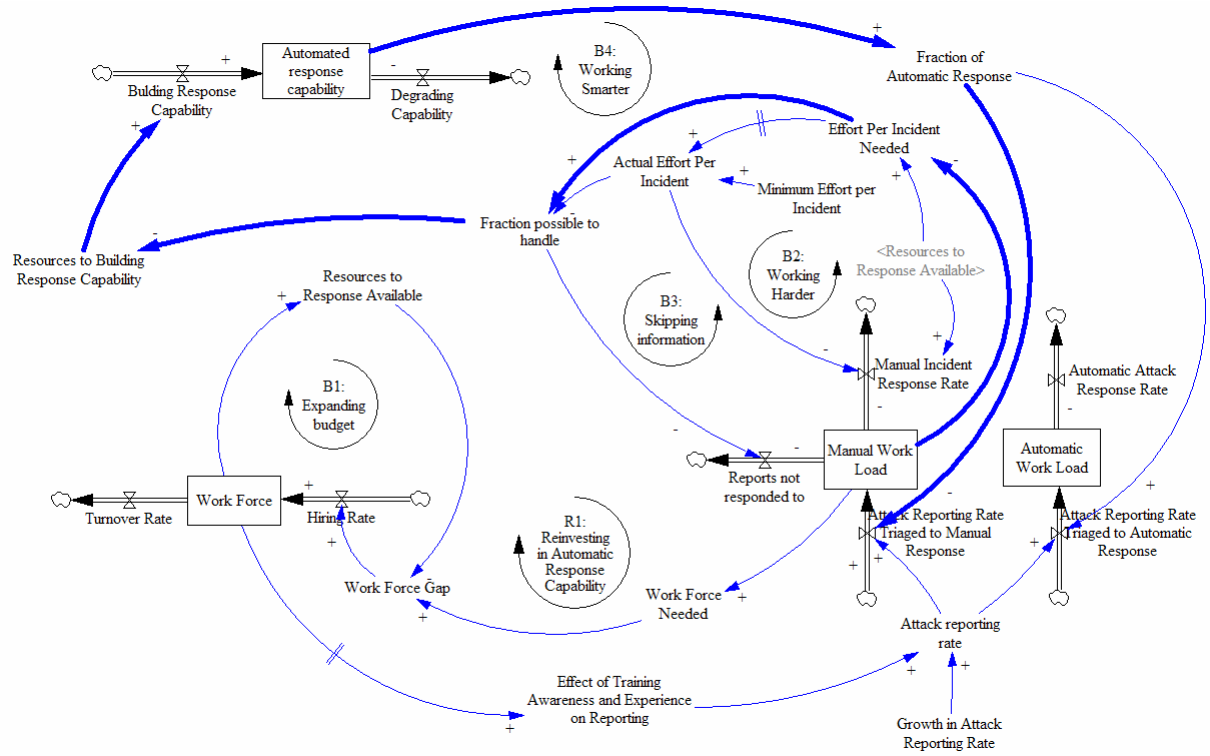


**Figure 10: Working smarter by investing in automated response capability**

As *Fraction Possible to Handle* starts to decrease, at some point management will realise that more fundamental solutions are needed. Hence, *Resources to Building Response Capability* is increased. Therefore, *Automated Response Capability* will increase, but with a delay: it takes time before any such system will be available. When it becomes available, a fraction or all of the reports can be responded to automatically (dependent on the development plan and ambition of the improvement project). In our case, we assume a fully automated system. We have added a new stock and flow structure to capture this process of building automation. As *Fraction of Automatic Response* increases, less reports will be handled manually. Consequently the investments in automated response will act to balance the growing work load. This identifies another balancing loop labelled “B4: Working Smarter” (Figure 10).

Working smarter would be preferable to working harder. However, our investigation indicates that this has not been achieved historically. Past efforts, initiated over the years to develop new tools in our case-study, have been unsuccessful, and most of them have been short-lived. The most notable example of such a failed project took place in 2002-2003, and consumed a lot of resources without yielding any results.

The model in Figure 10 can easily be considered a simple version of a CSIRT manager's mental model. If his or her decisions are influenced by an unawareness of some important compensating feedback that can counteract the improvement efforts, the situation in the CSIRT might even get worse.

### 6.3.5 Unintended Side-Effects and Misperception of Feedback

An important factor that prevents process improvement from being successful, and thereby limiting long-term enhanced effectiveness, is what Repenning and Sterman have labelled the "Capability Trap" (2002 p. 290). Enhancements to the response process take time to build; at the same time, there is a continuous pressure where the response staff has to prioritise between helping victims with their incidents, and developing tools to enhance the handling of incidents. A crisis on the customer side is actually the normal situation that is dealt with several times a day when a customer needs help mitigating an incident. Hence, the pressure to help victims is naturally high. With limited resources available staff will tend to postpone further tool development until a crisis has been resolved. This method of prioritisation can also be institutionalised to handle peaks in work load. In fact, it is described as common procedure in the CSIRT literature that resources should be prioritised to help victims when the CSIRT faces a heavy work load of incidents (West-Brown et al., 2003 p. 66). This becomes a significant problem if it gradually becomes the daily routine and not just a temporary solution during peaks in the work load.

Similar behaviour has been identified in other contexts as well, such as in production systems or in product innovation (Keating et al., 1999, Repenning and Sterman, 2001; Repenning and Sterman, 2002). The main reason that these authors attribute to such behaviour, is that people tend to focus on short-term tangible effects from cutting corners and working harder, as opposed to developing tools that will not be available until some time in the future. The results of such process improvement are much more intangible than if a staff member is able to help a victim right away. We might even consider this an even more pressing issue in a CSIRT context than in many other types of working environments due to the daily "crisis-environment" faced by a CSIRT response staff.

An interesting question, however, is why the CSIRT cannot buy such a commercial tool off-the-shelf instead of developing it in-house. The reason is that hardly any such commercial tools are available covering all the needs for a CSIRT (Killcrece et al., 2003 p. 127). Hence, the only option is to develop the tools internally in the CSIRT. As the incident response staff is normally not specialised in programming, the development of such tools may take longer than for skilled programmers. On the other hand, the incident handling experts are very likely to better understand the requirements needed for such tools, and they may therefore be better suited to specify the features than a specialised programmer. Gonzalez (2005) has argued that an important corollary of improving the performance of outsourced security-handling organisations, such as CSIRTs, should be a growing opportunity for development of professional and commercial tools.

As more resources are spent on building automated response capability, fewer resources will be available for response, and consequently, *Work Force Gap* will increase and create an even more pressing situation than before. As long as *Work Force Gap* is negative, meaning a negative pressure, or slack, the risk of compensating feedback undermining the improvement project of automatic response will be low. However, if the project has been started when most other solutions to balance the workload have already been used up, *Work Force Gap* will most likely be positive, and with some delay it will generate a significant pressure on the work force. This *Work Force Pressure* can quickly deplete the resources allocated to building new automated response capability and lead to reallocation of resources to response and thereby reduce *Work Force Gap* short term. Hence, loop "B5: Cutting Improvement Corners" is closed.

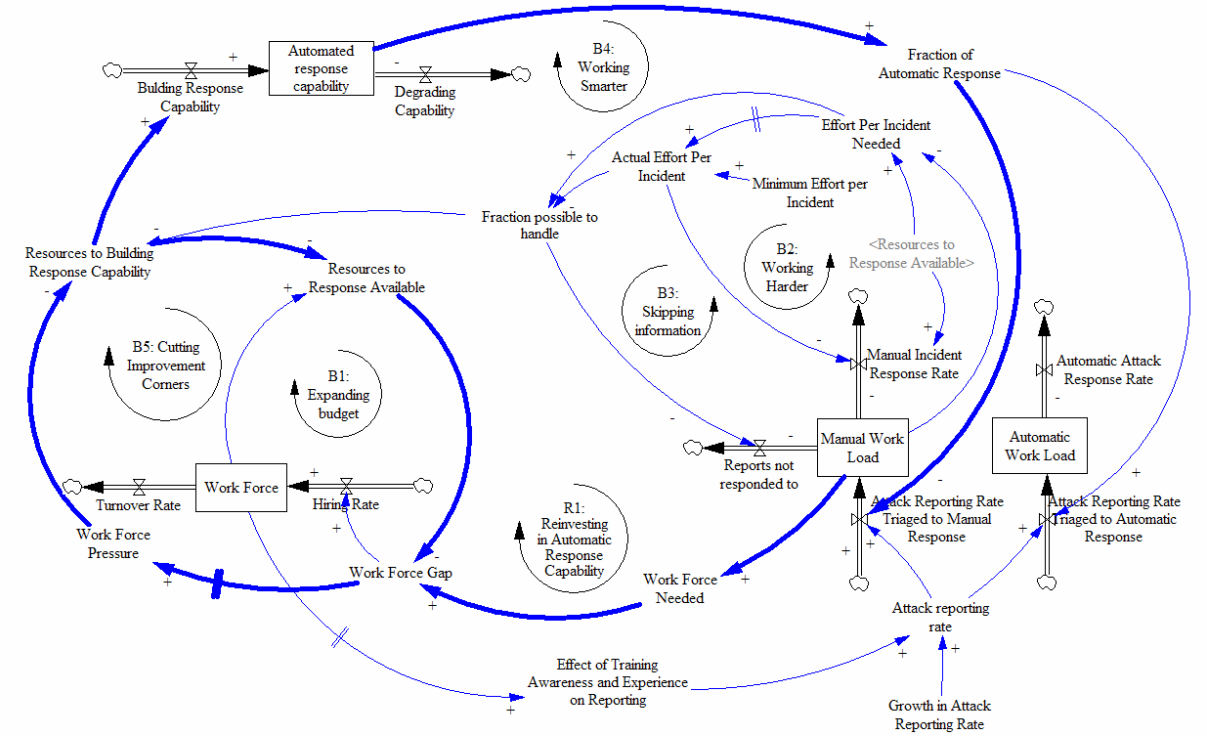


**Figure 11: Work force pressure can generate unintended side-effects**

According to Repenning and Sterman (2002), management may make the situation even worse by forcing people to work harder and harder to meet the desired response time. Bad performance, created by a vicious cycle of underinvestment in process improvement, may be attributed to poor worker mentality leading people to prioritise even more corner-cutting to meet the goals set by management. The real reason for such poor performance is the causal structure of the system that has been described here, and not the wrong attributions made by management. Such attribution errors of the causes to the problem can lead the system further and further away from the fundamental solution, and into a "Capability Trap".

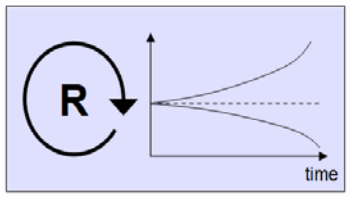


**Figure 12: A dominating reinforcing feedback loop will generate diverging behaviour of exponential growth or decline. Such a loop is therefore often referred to as a vicious cycle or virtuous cycle (dependent on the desire to grow or decline).**

If the balancing loop, "B5: Cutting Improvement Corners", is too strong and overwhelms the improvement process, the only thing

management has achieved is to make the situation worse by using resources that did not yield any results.

By closing the balancing loop B5, we have also created a new type of loop. “R1: Reinvesting in Automatic Response Capability” which is a reinforcing feedback loop that tends to amplify effects, good or bad, if it becomes dominant. For example, if the CSIRT develops more *Automated Response Capability*, *Manual Work Load* will decrease. Hence, as fewer resources are needed for response, *Work Force Pressure* decreases, and even more effort can be put into the improvement process generating a *virtuous* cycle of improvement. As one improvement project is finished, the new resources can then be allocated to even more improvement in other CSIRT processes as well, but that is outside the scope of the model presented here. Conversely, if *Work Force Pressure* is too high, it can lead to less development, more work consuming more resources, and hence, leads to even less improvement. In this case, the loop acts like a *vicious* cycle for the CSIRT.

## 6.4 Model Validation and Testing

There are several ways to validate the assumptions in a System Dynamics simulation model, but we will distinguish between two main types:

- Structural tests
- Behavioural tests

The structural tests generally mean that the structure is presented to experts, either through discussion or in graphical format, and they will then validate whether these assumptions correspond to the real world they know. Often, several different people will contribute to this process in a formal or informal way as it is seldom possible to find someone who knows everything about the problem and the system in question.

In the behavioural tests we typically evaluate the model output, for example by comparing with historical output. Any difference between the two outputs will then be discussed to understand the difference. Often this process uncovers incorrect assumptions in the structure of the model. Obviously, the main point is to make the model generate the right output for the right reason.

In Figure 13, we describe some examples of historical data compared to the model output. The key variables represent the work load, as well as the resources available to handle the work load. In other words, the main resources have to be balanced for the CSIRT to stay effective. The lines in grey reflect the historical figures, while the red lines represent the corresponding variables in the simulation model.

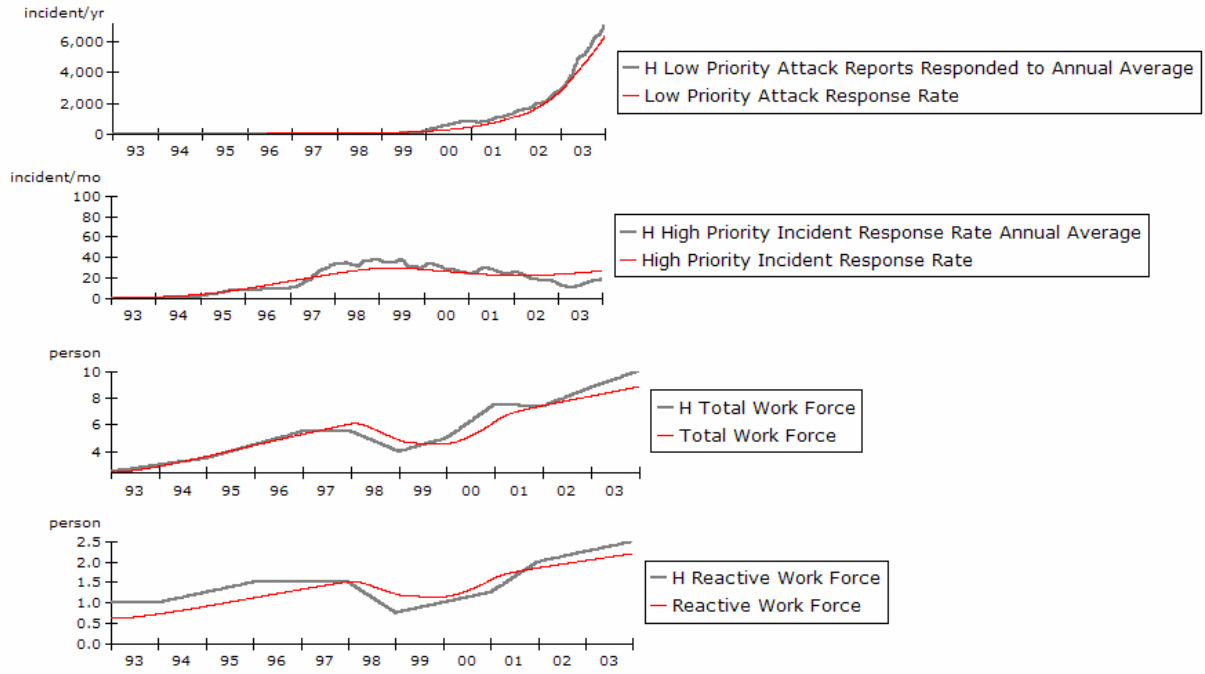


**Figure 13: Examples of model output compared to historical data**

The data available for comparison has been very scarce, and in some cases we have also faced problems with changing definition of data over time. For example, what has been considered an incident and its priority has changed over time. Reference modes for other variables have been based on more anecdotal information.

In the following discussion of anecdotal information, it would be interesting to see how the model replicates history with respect to the compensating feedback of working harder, and skipping information. In addition, we would like to see how the model also creates an environment where improvement efforts are undermined. Next, we will discuss some of the main results and corresponding assumptions in the model (see Figure 14).

Prior to 1999, the CSIRT compensated for any growth in work load by working harder. Hence, we see that *Actual Effort Per Low Priority Incident* gradually declines, indicating that loop "B2: Working Harder" is the main compensating factor that keeps the system in balance. At the beginning of 2002, we see an emerging gap between *Low Priority Attacks Reporting Rate* and *Low Priority Attacks Handling Rate*. This indicates that loop "B3: Skipping Information", gradually increases in importance. At the same time, the problems of handling the work load gradually also increase *Work Load Gap Pressure*, and at almost the same time, a project for developing *Automated Response Capability* is initiated as *Resources Allocated to Response Improvement* starts to grow. However, this variable is never matched by *Effort to Response Improvement* due to *Work load Gap Pressure*. During 2003, this pressure becomes so large that even though more resources are allocated to process improvement, the effort declines and the project is discontinued. As a matter of fact, such a project was indeed taking place in 2002-2003, and according to one of the participants, it was referred to as a "Nightmare" due to the constant high work pressure. It was also mentioned that the project was too ambitious and unfocused to be successful.

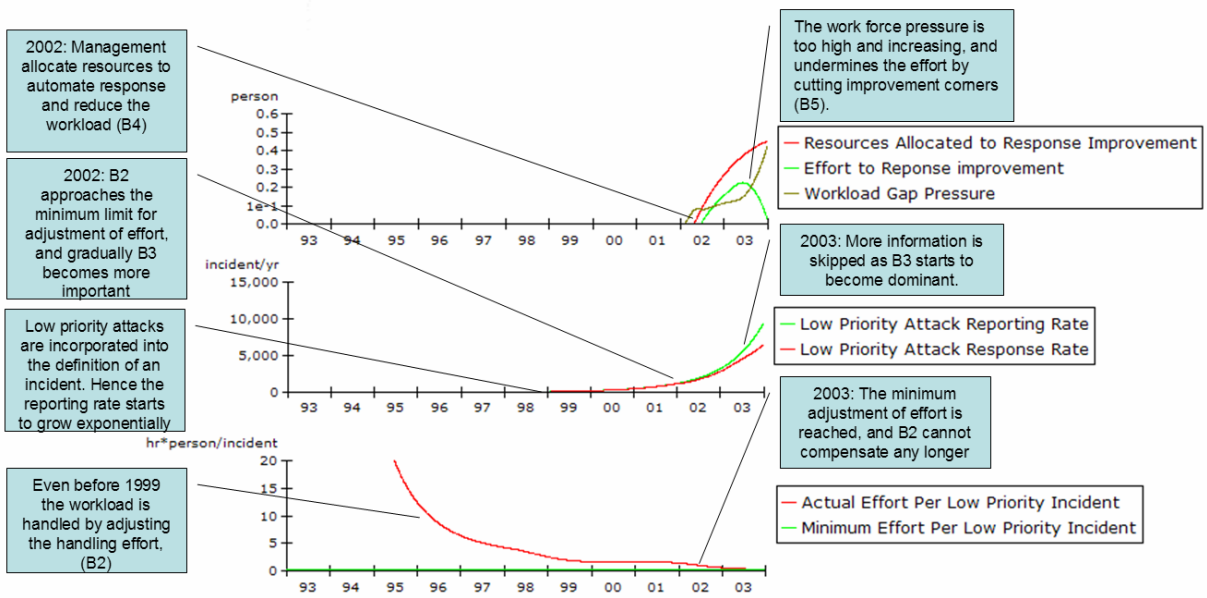


**Figure 14: Simulation results discussed based on anecdotal information about past improvement failures.**

At the end of this quick validation of the model, we would also like to mention yet another common behavioural test. This is to give the model an extreme input. As a matter of fact, such an input is very much a part of the problem definition in our model, namely, the exponential growth in low priority attack reports that has doubled every year since 1999. In particular, this extreme input leads us to ask questions where we gradually were able to understand the balancing feedback of loop B2 and B3, and build it into the model.

We assume the simulation model produces a very good output if we compare it both to time-series of data, as well as more anecdotal information. Hence, we conclude that the assumptions in the model are strengthened by the fact that the outputs closely resemble the real-world history of the CSIRT. Validation is really a never-ending story, but through testing, we will gradually build confidence in the simulation model, and get a better understanding of our complex dynamic problem.

## *6.5 Policy Experimentation and Future Scenarios*

When a problem is well-understood, we can try to solve it. This is typically accomplished by altering some of the policies in the system, or by reengineering parts of it.

What can a CSIRT do to improve the situation? One possible step in the right direction, given by Keating et al. (1999), is to increase the slack for people. Such an approach will ease the pressure for short-term symptomatic solutions. This can be done in two ways:

1. by adding resources, or
2. reducing the work load

The implementation of a solution may vary between types of CSIRTs, for example, in the way they are getting their funding, how this is linked to their services, whether it is an internal CSIRT in a larger organisation, or whether it is an external CSIRT with limited authority in their constituency. Normally, it is not easy to get funding for more people in an already overstretched budget. Increased funding is often linked to establishing new services, something that would add to the existing workload even more. Consequently, the first option mentioned above (adding resources) might not be feasible. Making matters even worse, the

information overload created by the increasing rate of attacks reported, will most likely just move the system even further away from a solution. Even if more resources are added, they will most likely not compensate enough for the exponential growth.

This leaves management with the second option, namely to reduce *Manual Work Load*. One way to do this is to change the scope of service provided by the CSIRT. This adjustment of *Current Scope of Service* to reduce *Manual Work Load* is shown in Figure 15.

As less of the information reported is responded to, *Fraction Possible to Handle* decreases. If this fraction decreases below a certain point, the service for responding to low priority attack reports may be temporarily discontinued. Hence, certain reports are removed from *Manual Work Load* through the outflow *Reports not Responded to*. This effect has been modelled as a balancing feedback-loop, "B6: Adjusting Scope of Service". This loop can be seen as a formalisation of the balancing loop "B3: Skipping Information", but it goes beyond that by reducing the work load so much that the necessary resources are released for other purposes. All high priority incidents are, of course, handled as usual, but that process is not discussed here in further detail.

We tested the model with the following assumptions:

- the low priority service was discontinued when a certain fraction of the reports were not handled
- the low priority service was reinitiated when the automated response capability reached 100%. That is, a fully automated response system for low priority incidents was in place.

Finally, we extended the simulation time by 3 years to see the long-term effects.

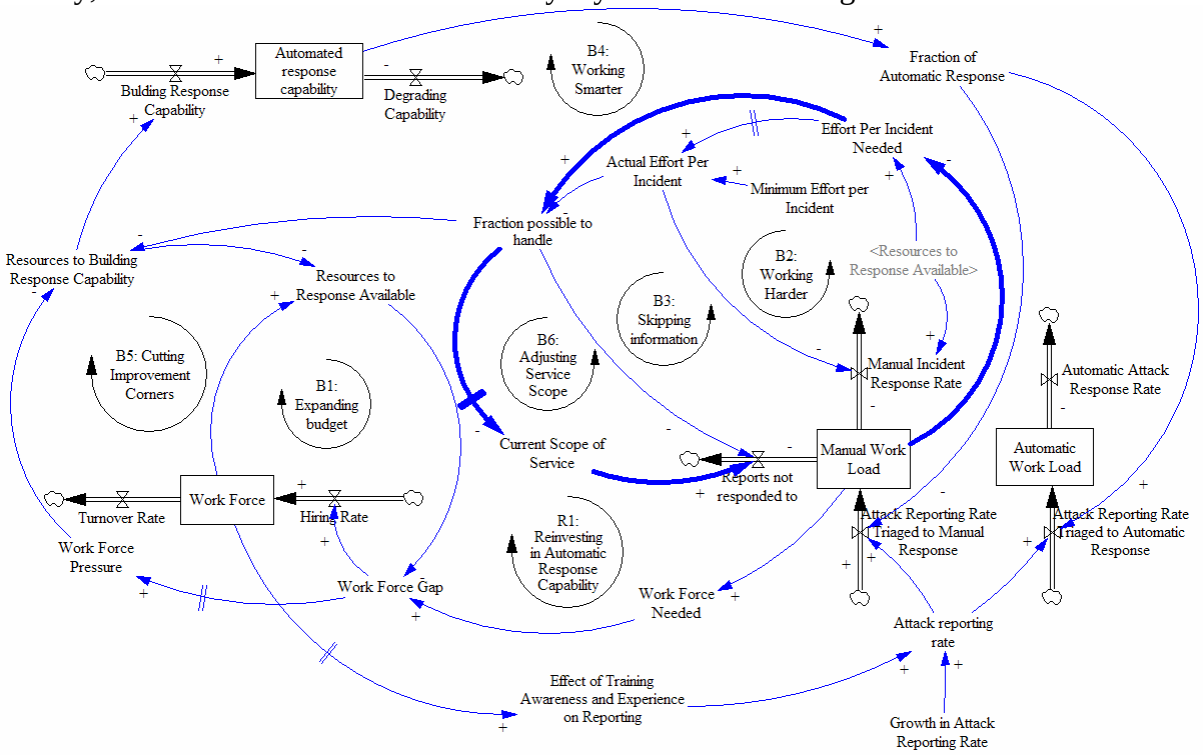


**Figure 15: Adjusting service scope to free up resources**

The results of these changes show that the suggested strategy is a worse-before-better solution as can be seen in Figure 16.

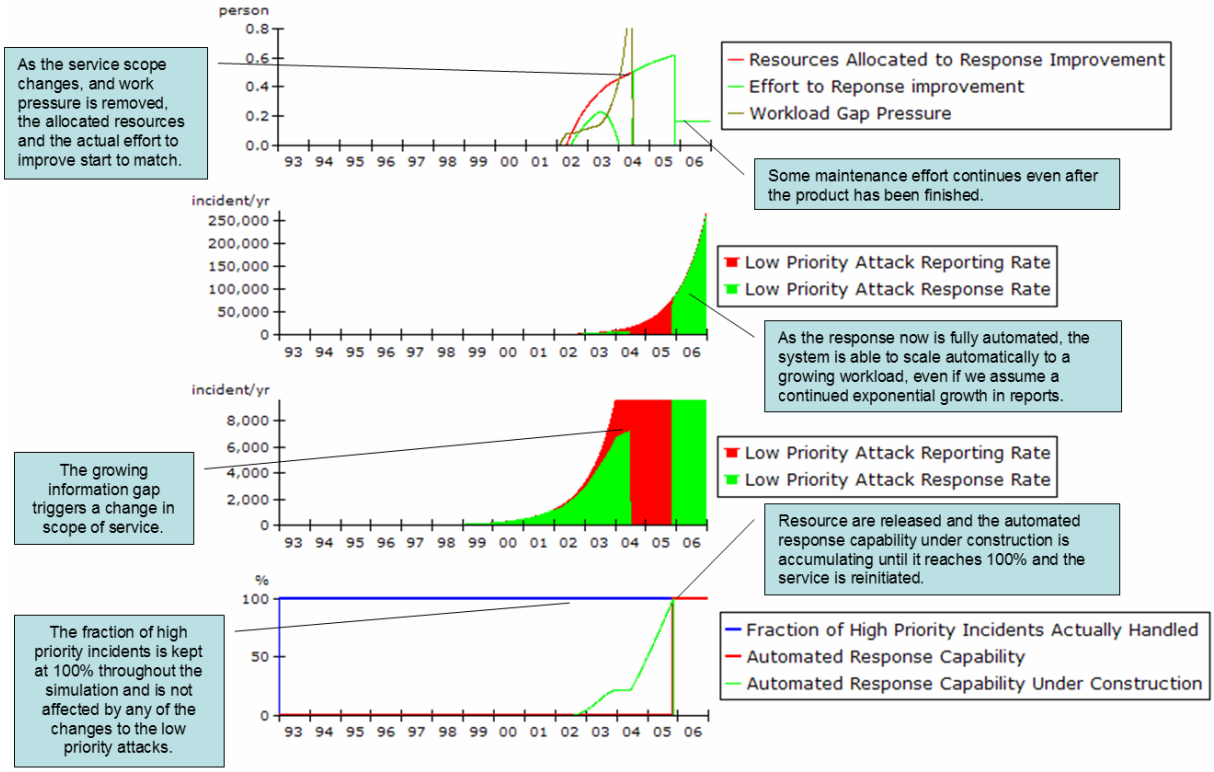


**Figure 16: Future scenario - Temporarily changing service scope**

In the middle of 2004, *Fraction Possible to Handle* decreases beyond a certain threshold set by management. Hence, low priority incidents are excluded from the scope of service. This makes *Work load Gap Pressure* drop immediately. Hence, *Resources to Building Response Capability* quickly increase to the level of *Resources Allocated to Response Improvement*. The development project can thereby continue until it is completed. That is, when 100% of all low priority attacks can be responded to automatically. At this point in time the scope of service is increased again back to its previous level, and all attacks reported are responded to. Due to the automated response, the system is able to scale automatically to the work load as well.

Just to avoid any misunderstanding of the incident handling process, we have also observed that the fraction of High Priority Incidents Actually Handled stays at 100% throughout the simulation. Hence, the core function of the CSIRT is not notably influenced by the scenario.

Based on the assumptions in the model, the scenario indicates that it is difficult, but not impossible, to escape the capability trap. The solution we outline here is a classical example of a worse-before-better scenario. However, there are certain issues that should also be mentioned before this strategy is considered for implementation.

A CSIRT implementing an automated response capability might become so efficient that it will generate information overload for other CSIRTs or constituents in a similar way that the CSIRT we have investigated have experienced in the past. Hence, the CSIRT must make sure that the receiver is not overloaded with information. A potential solution is to use a pull, instead of a push system. Instead of pushing out responses to constituents, the CSIRT can post a recommended response on a web page, and just notify the receiver of any updates. Hence, it is up to the receiver to check out the postings. Another concern is that unless response is made

in a careful way, it may be considered of less value by the receiver than if it seems to be manual. Automatic response falls into the danger of being considered “spam”.

If the CSIRT expands its *Current Service Scope* even further after the situation has improved, the constituency will, with a delay, start reporting more as they might even automate how they submit such information to the CSIRT. This means that new or expanded services can easily eat up any slack in the system, and make process improvement even more difficult unless there is enough capability in the form of automated response tools available to handle the increasing work load. Consequently, it is important that management is not tempted to reduce the slack so much that they prevent further improvements from taking place after the response capability has been enhanced.

# 7 Future Research

As previously mentioned, and exemplified through the study of one particular external CSIRT, the “Capability Trap” seems like a very applicable description of the problem faced by many CSIRTs today. Given the limitations of our inductive approach, further research is needed to support these preliminary findings.

In this paper we have focused on enhancing responsive services, very often referred to as reactive services in a CSIRT (West-Brown et al., 2003). We have also found indications that the “Capability Trap” can be a useful description for limits to process improvement of other services as well, such as proactive services. We can assume there are interactions between proactive services, reactive services, and detection. An iceberg is an illustrative example: Reactive services handle what the CSIRT can see of the iceberg, proactive services will reduce the size of the ice-berg, and detection systems can make more of the iceberg visible to the CSIRT. By expanding our analysis beyond just process improvement of reactive services, the picture will become increasingly more complex. As complexity increases, process improvement goals will also become increasingly more difficult to achieve (Keating et al., 1999). For example, instead of focussing on response capability improvements by developing new tools, the CSIRT can try to lower the work load by proactively trying to educate their constituency, run penetration tests, issue advisories, etc. However, this means crossing multiple organisational boundaries within the constituency to reach out to thousands, and maybe even millions, of individual users. In addition, more proactive efforts will increase the awareness about the CSIRT, which might lead to even more incidents being reported, and correspondingly, increasing, and not decreasing, the responsive workload. To proactively avoid incidents, more interaction and information exchange regarding vulnerabilities, solutions, workarounds etc., between vendors, other CSIRTs and other organisations, will also be necessary to make the impact of proactive efforts more effective. Many such initiatives are underway in the CSIRT community, but it will most likely take time before we can see the real effect of these efforts. While the “Capability Trap” can be applicable to other interactions within a CSIRT, any suggested solutions to escape the “Capability Trap”, in the context of process improvement of the response service, may not be transferred to other contexts so easily. We would expect that it will be increasingly difficult to achieve process improvement as the complexity of interacting services increases.

# 8 Conclusion

The inductive approach of this research means that we cannot easily generalise beyond the particular CSIRT case we have studied, although we expect that the findings are interesting for other coordinating CSIRTs, in particular. Our preliminary results indicate that the theory

of the “Capability Trap” is useful for understanding why a CSIRT can experience problems improving to stay effective. The typical over-stretched resource situation in a CSIRT that limits its effectiveness can lead it into the “Capability Trap” that forces the CSIRT to work harder and harder which further reduces its capability to improve. People under pressure will tend to choose short-term solutions, such as working harder, to achieve tangible results quickly. This is opposed to long-term solutions, such as tool development, where the effect is delayed and more intangible. A CSIRT that has over-stretched its resources over a long time period must be prepared to go through a worse-before-better scenario to escape the “Capability Trap”. Such a transition process can be quite painful to the CSIRT and its surrounding environment, for example, through adjustments to scope of service to release resources for improvement.

Due to its constantly changing environment, a CSIRT must always ensure it has enough resources available to be able to continuously improve the capability and quality of its processes and services by reinvesting resources made available from previous improvements into future improvements, and consequently creating a virtuous reinforcing cycle of sustained effectiveness.

## Acknowledgements

This work would never have been possible without the help and aid from DFN-CERT Services GmbH. Certain people have made a significant contribution to the work presented in this paper. In no particular order of importance, Andreas Bunten and Marco Thorbrügge currently working in DFN-CERT Services GmbH have provided substantial information for the last couple of years of operation until the present day. Peter Janitz (former employee in DFN-CERT, now with SECUNET Security Networks, Germany) gave us insights into the transition period when DFN-CERT was transformed from a research project into a non-profit company in 1999. Wolfgang Ley (former employee in DFN-CERT, now with SUN Microsystems, Germany) helped us with an in-depth understanding of the history of DFN-CERT prior to the transition in 1999. They were all contributing to our understanding, and thereby our ability to create the holistic theory in the form of a simulation model that we have presented in this paper.

In addition to the funding provided by Agder University College in terms of a PhD-fellowship and travel assistance, we acknowledge support from the Strategisk Høgskoleprogram in Mobile Communication and from IKTSoS, the research program for ICT Security and Vulnerability (both funded by the Research Council of Norway).

# Literature

Blakley, Bob (1997): "The Emperor's Old Armor," *Proceedings of the 1996 New Security Paradigms Workshop*, Lake Arrowhead, California, September 17-20, 1996, Association for Computing Machinery.

Forrester, Jay W. (1994): "Policies, Decisions, and Information Sources for Modeling" p. 51-84 in Modeling for Learning Organizations, John D.W Morecroft and John D. Sterman (ed.), Productivity Press, Portland Oregon.

Fisher, D.A. and Lipson H.F. (1999): "Emergent Algorithms — A New Method for Enhancing Survivability in Unbounded Systems," *Proceedings of the 32nd Annual Hawaii International Conference on System Sciences*, Maui, Hawaii, January 5-8, (HICSS-32), IEEE Computer Society.

Gonzalez, José. J.(2005): Towards a Cyber Security Reporting System – A Quality Improvement Process. Proceedings to The International Conference on Computer Safety, Reliability and Security – SAFECOMP 2005, Fredrikstad, Norway.

Keating, Elizabeth K.; Oliva, Rogelio; Repenning, Nelson P.; Rockart, Scott and Sterman, John D. (1999): Overcoming the Improvement Paradox: European Management Journal, Vol. 17, No. 2, pp. 120-134, 1999.

Lipson, Howard and Fisher, David A.(1999): "Survivability – A New Technical and Business Perspective on Security" Proceedings of the 1999 New Security Paradigms Workshop. Caledon Hills, ON, September 21-24, 1999. New York, NY, Association for Computing Machinery.

Lipson, Howard (2002): Tracking and Tracing Cyber-Attacks: Technical Challenges and Global Policy Issues. SPECIAL REPORT CMU/SEI-2002-SR-009, Pittsburgh, PA, USA.

Killcrece, Georgia; Kossakowski, Klaus-Peter; Ruefle, Robin and Zajicek Mark (2003a): State of the Practice of Computer Security Incident Response Teams (CSIRTs). Technical Report. CMU/SEI-2003-TR-001, Pittsburgh, PA, USA.

Killcrece, Georgia; Kossakowski, Klaus-Peter; Ruefle, Robin and Zajicek Mark (2003b): Organizational Model for Computer Security Incident Response Teams. Handbook. CMU/SEI-2003-HB-001, Pittsburgh, PA, USA.

Kossakowski, Klaus-Peter (2000): Information Technology Incident Response Capabilities, Libri Books on Demand, Hamburg, Germany.

Repenning, Nelson P. and Sterman John D. (2001): Nobody Ever Gets Credit for Fixing Problems that Never Happened - Creating and Sustaining Process Improvement. California Management Review Vol 43. No.4 Summer 2001.

Repenning, Nelson P. and Sterman John D. (2002): Capability Traps and Self-Confirming Attribution Errors in the Dynamics of Process Improvement, in Administrative Science Quarterly, 47, page 265-295.

Sawicka, Agata, Gonzalez, Jose J. and Qian, Ying (2005): Managing CSIRT Capacity as a Renewable Resource Management Challenge: An Experimental Study. This Proceedings.

Smith, Danny (1994): Forming an Incident Response Team, Proceedings of the FIRST Annual Conference. University of Queensland, Brisbane, Australia, July 1994.

Sterman, John D.( 2000): "Business Dynamics – Systems Thinking and Modeling for a Complex World", Irwin McGraw-Hill.

Sterman, John D.(2002): All models are wrong: reflections on becoming a system scientist. System Dynamic Review vol. 18 no. 4.

van Wyk, Kenneth R. and Forno, Richard (2001): Incident Response, O'Reilly and Associates, Inc., Sebastopol, CA, USA.

West-Brown, Moria J.; Stikvoort, Don; Kossakowski, Klaus-Peter; Killcrece, Gerogia; Ruefle, Robin and Zajicek Mark (2003). Handbook of Computer Security Incident Response Teams (CSIRTs) Second Edition (CMU/SEI-2003-HB-002), Pittsburgh, PA, USA.

Wiik, Johannes and Kossakowski, Klaus-Peter (2005): Dynamics of Incident Response. 17th Annual FIRST Conference on Computer Security Incident Handling June 26–July 01, 2005 — Singapore Hosted by FIRST.Org, Inc.

# Appendix: Full Overview of Model