Jens Nedon, Sandra Frings, Oliver Göbel (Hrsg.)

IT-Incident Management & IT-Forensics

Erste Tagung der Fachgruppe SIDAR der Gesellschaft für Informatik

24. – 25. November 2003
in Stuttgart, Deutschland

Gesellschaft für Informatik 2003

Lecture Notes in Informatics (LNI) - Proceedings
Series of the Gesellschaft für Informatik (GI)

Volume P-39

ISBN 3-88579-368-7
ISSN 1617-5468

Volume Editors
Jens Nedon
ConSecur GmbH
Schulze-Delitzsch-Strasse 2, D-49716 Meppen
Nedon@consecur.de
Sandra Frings
Fraunhofer-Institut für Arbeitswirtschaft und Organisation IAO
Nobelstraße 12, D-70569 Stuttgart
Sandra.Frings@iao.fhg.de
Oliver Göbel
RUS-CERT (Universität Stuttgart)
Breitscheidstr. 2, D-70174 Stuttgart
Goebel@CERT.Uni-Stuttgart.DE

printed by Köllen Druck+Verlag GmbH, Bonn

Visual Problem-Solving Support for New Event Triage in Centralized Network Security Monitoring: Challenges, Tools and Benefits

Markus Stolze, René Pawlitzek, Andreas Wespi

IBM Research Zurich Laboratory
Säumerstrasse 4
CH-8803 Rüschlikon
{mrs, rpa, anw}@zurich.ibm.com

Abstract: Organizations that provide centralized security monitoring of the networks of multiple third-party organizations are faced with a challenging task. The amount of security event data to be processed presents not only a technical challenge, but also a problem-solving challenge to operators. We present a model of the problem-solving process and discuss how visual support tools can facilitate the central problem-solving step called new event triage. We argue that with tools such as these the natural benefits of centralized monitoring can come into play, which enhances effectiveness of centralized monitoring to a level beyond the reach of organizations focusing exclusively on their own network.

1 Introduction

Today more and more organizations are outsourcing network security monitoring to trusted organizations that run centralized managed security operations centers (SOC) performing rapid-response round-the-clock monitoring of multiple customer networks in parallel. Running such a SOC is a challenging task because the data arriving from a diverse set of intrusion detection systems (IDS) from the customer sites has to be integrated and processed. Even a single signature-based network IDS will generate a large number of security events if it is installed in a network that has a high level of “noise” due to misconfigured services, user accidents, damaged/lost data packets, network management services, heartbeat information, and other activities unrelated to intrusion attempts [HWG 02]. This noise level contributes to a high false-positive alert rate, i.e. the reporting of network activities that match signatures, but are not part of an attack. In such situations operators see a large number of security events, most of which, in most situations, can be safely ignored.

Recently we studied the tasks of operators working in a SOC that provides a 24/7 network security monitoring service to multiple customers [SPH 03]. The operators we observed work in shifts, with multiple operators per shift. The main task of the operators is to identify and report to the respective customers any abnormal network activity detected by the IDSs installed in the customer’s network. In that setup we observed that primarily signature-based IDS are monitored. Each of the monitored IDS generates an average of ten security events per hour with large variances between IDSs and over time. Operators take turns monitoring the events dropping in from the more than 500 IDSs. If a suspicious-looking event or pattern of events is observed a “call-out” procedure is started. The goal of this procedure is to ensure that the person responsible at the affected customer’s site is made aware of the situation and is provided with suggestions on how to deal with the observed problem.

Monitoring of IDS events is only one of the tasks performed by the operators in the team. The team also has to monitor the status of the deployed IDSs, stay abreast of information from vendor-specific and general security sources, log their main activities, and compile cumulative reports for customers.

Setting up and operating the technical infrastructure for such a SOC is a complex task. For example, the infrastructure has to collect securely and reliably in real-time the flow of security events arriving from the heterogeneous set of IDSs. We will not discuss here how to best address these technical challenges. Instead, our focus is on the challenges that individual operators face while monitoring the rich stream of IDS events. Below we first describe the problem operators face and derive requirements for support tools. We then present a set of information visualization techniques we developed and review related research on problem-solving support tools for security monitoring. We then conclude that with the help of such tools, the natural benefits of a centralized service can come into play, which enhances the effectiveness of centralized monitoring to a level beyond the reach of organizations focusing exclusively on their own network.

2 SOC Operators’ Problem-Solving Task

The problem to be solved by operators monitoring a stream of IDS events is to determine when and how to inform a customer about a possibly problematic situation in their network. The difficulty is that all problematic situations should be communicated, but no unnecessary communication should occur. This is not an easy task because IDSs, especially in “noisy” networks, produce a large number of false alarms, i.e. events that do not indicate an actual security problem. The task is further complicated by the fact that IDS events are characterized by a large number of properties. We observed that five of these properties are most frequently used for the initial event triage: Event severity, alarm type, attack origin, attack target, and attack signature. In the more detailed analysis of special cases, operators investigate additional properties more closely as well as collect and evaluate external information. Before contacting a customer, operators perform a lookup in a log database to determine whether the observed problem has been reported previously.

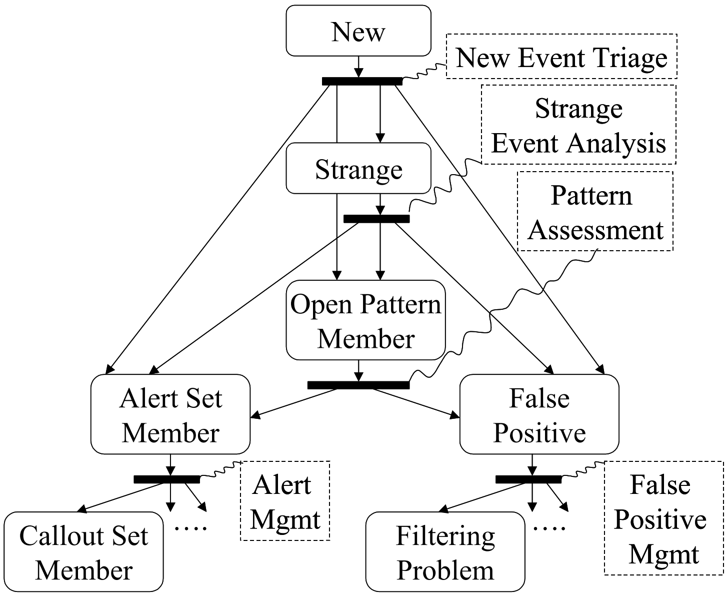


Figure 2.1: State transition diagram describing the multilevel event classification process of newly arrived events.

We have conceptualized the above description of the SOC operators' task in the following descriptive model (cf. Figure 2.1). According to this model operators go through multiple levels of event classification to process the incoming stream of new security events. The stream consists of base events generated, for example, by an IDS. It can also contain summary events generated, for example, by a correlation infrastructure. Whereas the model focuses on the state transitions of a single event it should be noted that operators frequently act on multiple events in parallel. Operators initially perform a quick *new event triage* in which they decide whether a new event is obviously part of a sequence of events that a customer needs to be alerted about (classification "alert set member"), whether an event requires more detailed analysis (classification "strange" event), or whether an event is an obvious false positive. Furthermore, the operator can also decide that the newly arrived event is obviously part of an open pattern of events (classification "open pattern member"). These events then need to be further monitored before operators can assess whether the events in the pattern should be classified as "alert set members" or "false positives". As operators will often process events in sets, part of performing the new event triage is to determine which set of events is the most interesting or the most urgent. They must then identify the main characteristics of the events in the set and make sure that the set is homogeneous enough to justify batch treatment. If necessary, operators have to focus on a subset of events to ensure homogeneity.

In the next step, *strange event analysis*, operators inspect the properties of strange events or sets of strange events in more detail. As part of this, operators collect and consider additional context information that was not automatically added to the event properties. For example, in cases in which operators want to know more details about an IP address they can collect this information from an external "who is" service. The result of this event analysis is that the analyzed event (or set of events) is classified as being either a member of an alert set, a false positive, or a member of an open pattern.

While performing new event triage and strange event analysis, operators also have to monitor the open event patterns to quickly identify situations in which an open pattern is completed and a customer needs to be alerted (*pattern assessment*). An open event pattern is a sequence of events for which the operator has decided that it makes sense not to deal with this sequence immediately, but to wait for additional events before making a decision. In some cases the member events of an open pattern can be classified as false positive. This is usually the case after a pattern has timed-out or completed in a way that proved to be harmless.

Once operators have identified a set of events that warrants alerting the customer, they also have to decide how to deal with the situation (*alert management*). The decision can be, for example, that the situation is urgent enough (or new enough) to justify calling the customer on the phone. Operators might also decide that an e-mail report is sufficient or that the information can be sent as an appendix to an already sent report. For the events that are classified as false positive, operators have to decide whether they can take action and modify the infrastructure setup in a way that they do not need to deal manually with similar events in the future (*false positive management*).

3 SOC Problem-Solving Support

Tools that support operators' problem-solving tasks should facilitate the steps mentioned above. In particular, these tools should provide the following problem-solving support functionality.

3.1 New Event Triage Support

Support tools should make it easy for operators to quickly scan newly arrived events, be alerted about interesting patterns in sensor activity, and compare current sensor activity with recent sensor activity. Tools should proactively collect and correlate event context information. The results of this should be used to quickly bring to the attention of operators the most important events and event properties. For example, when looking at an event it is often useful for operators to know whether the identified source of a given problematic network traffic is a computer owned by the customer organization.

Tools should also support operators in gaining and maintaining "situation awareness". That is, operators should be informed about all important aspects of the current problem-solving situation. Tools therefore should make it easy for operators to stay abreast of new types of exploits and vulnerabilities, recent changes in the infrastructure, as well as sensor and customer-specific information such as scheduled maintenance work and security scans.

3.2 Strange Event Analysis Support

Ideally operators will have to deal with very few strange events because the problem-solving support tools enable them to deal quickly with most events during the initial new event triage step. To facilitate event analysis, tools should make it easy for operators to reach third-party information sources and integrate and compare the collected information. For example, integration of “who is” information of a source IP and display of vulnerability check and background information on a target IP should be provided. Tools should also facilitate saving, marking, and forwarding strange events to other operators.

3.3 Pattern Assessment Support

Support tools should make it easy for operators to monitor the arrival of events that complete an already identified partial event pattern. Tools should also help operators to manage open patterns over time and to remove outdated patterns.

3.4 Alert Management

Tools should make it easy for operators to determine quickly how urgent it is to inform the respective customer of an alert situation. Tools should also support the efficient and reliable execution of the selected form of communication with the customer.

3.5 False Alarm Management

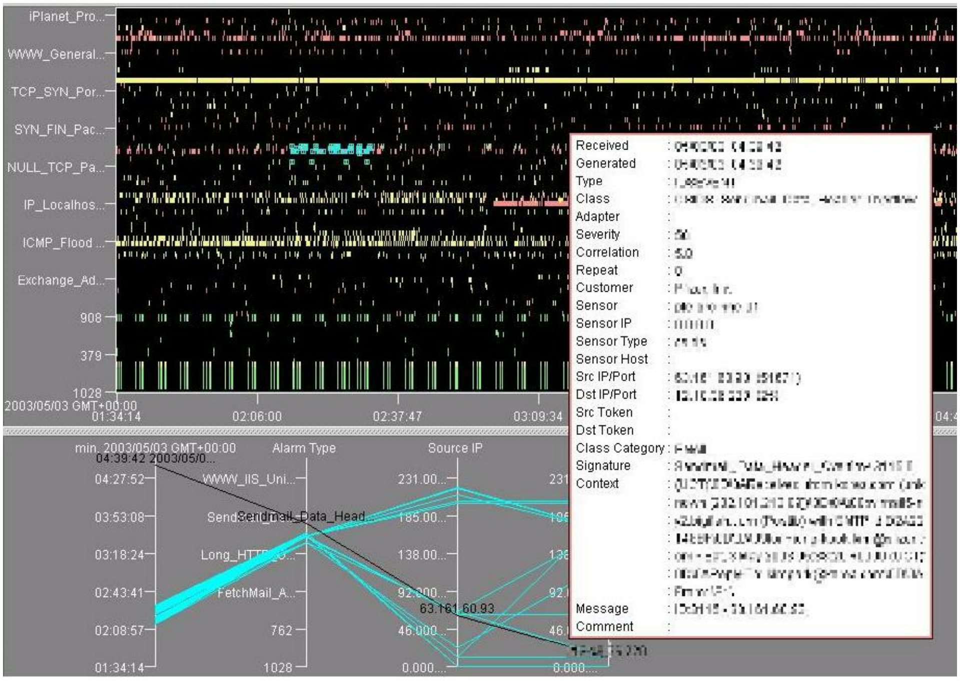
Tools should make it easy for operators to inspect why the identified security events were generated and provide means to debug, probe and adapt the infrastructure. For example, operators should be provided with means to monitor and adapt the rules that automatically filter out unproblematic events.

In this paper we discuss how visualization of security event data can support the task of new event triage. Of course, to provide a comprehensive solution and enable operators to perform their monitoring tasks in an optimal way, a system has to support all problem-solving steps.

4 Visual Support for New Event Triage

In our recent research we have experimented with and developed interactive information visualizations of a security event stream. The visualizations are specifically geared towards supporting the task of new event triage. Our current visualization combines a scatter-plot and parallel-coordinates display of the event data (cf. Figure 4.1).

The x-axis of the scatter-plot is fixed to be the generation time of events. Operators can choose the y-axis to map any other of the main event properties such as event type, severity, attack origin, and attack target. The visualization also includes a parallel-coordinates display of events at the bottom of the window. A parallel-coordinates display is a specialized presentation of multidimensional data points. It typically shows a vertical coordinate system for each data dimension and displays individual data points as polygons connecting the values in each dimension [ID 99]. The scatter-plot and the parallel coordinate display are connected as coordinated views. Events selected in the scatter-plot (e.g. by rubber-banding) are displayed as lines in the parallel-coordinates view. This combination of visualization techniques facilitates the rapid inspection and comparison of events in a selected set of events. Figure 4.1 gives an impression of a real stream of event data visualized in such a way.



Received : [illegible]
Generated : [illegible]
Type : [illegible]
Class : [illegible]
Adapter :
Severity : [illegible]
Correlation : 5.0
Repeat : 0
Customer : [illegible]
Sensor : [illegible]
Sensor IP : [illegible]
Sensor Type : [illegible]
Sensor Host :
Src IP/Port : [illegible]
Dst IP/Port : [illegible]
Src Token :
Dst Token :
Class Category : [illegible]
Signature : [illegible]
Context : [illegible]
Message : [illegible]
Comment :

Figure 4.1: Coordinated scatter-plot and parallel-coordinates display of a security event stream (some data has been anonymized for privacy reasons).

We currently investigate ways in which this basic visualization can be enhanced to further facilitate the triage of new events. One example of such an enhancement is the “overlaid” multicoordinate scatter-plot display (cf. Figure 4.2). The x-axis still represents the generation time of an event and the y-axis a selectable event property. The special feature of this augmented scatter-plot is that the system performs a background scanning for simple event patterns and brings these patterns to the attention of the operators by providing a “shine through” display of the identified event patterns. This way the system is able to alert operators of event patterns that occur in any of the major event dimensions. Without such support these patterns might only be detected after a delay, and sometimes the existence of such a pattern in less frequently monitored event dimensions might go unnoticed. Such a visualization will not only reduce the time needed to identify an interesting set of events for immediate evaluation, it could also increase the efficiency of operators because the existence of a pattern often indicates that the events forming pattern can be classified together to be either alerts or false positive.

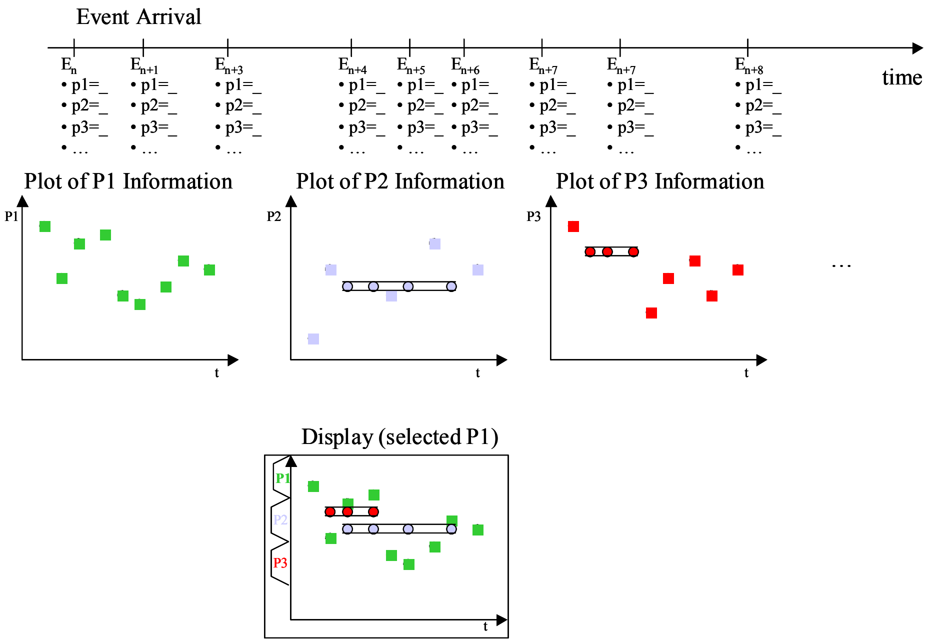


Figure 4.2: The multicoordinate scatter-plot visualization of an event stream as shown at the bottom.The figure is generated by overlaying scatter-plots from the different event dimensions and only allowing identified patterns to "shine through".

Although we have not yet performed formal user studies with the developed real-time information visualization, we have anecdotal reports from operators using the visualization in the context of a monitoring console for intrusion events. We developed this console as a Java-based add-on component for the IBM Tivoli Risk Manager security information management infrastructure.

The reports provide evidence that operators benefit from the information visualization by using the scatter-plot display for a rapid overview of the stream of incoming event data. They use it to identify particularly frequently occurring values, infrequently occurring values, recent changes in value frequency, and regular patterns in the occurrence of events with a specific value in the viewed dimension. This way, operators are able to gain situation awareness and to identify in the scatter plot the sets of events they need to evaluate next. Operators then select the identified set of events in the scatter plot and use the resulting display of events as lines in the parallel-coordinates view to identify the most important properties of the selected events and to ensure that the set is homogeneous enough to allow joint treatment of the involved events.

The reports also indicate that, with the provided support, operators are able to manage the large volume of event data that arrives from the multiple sensors. Of course this benefit cannot be attributed exclusively to the visualization of event information that is focused on supporting the triage of new events. Apart from the information visualization the console provides a number of other important functions including a tabular display of event information, support for rapidly focusing on the output of individual sensors, integration of event context information from third-party tools, and support for explicit event labeling and annotation. The other support functions are most likely equally important and in some cases even prerequisites for the visualization. Still, it is crucial for the effectiveness and efficiency of the overall monitoring task to deal quickly with events during a new event triage without having to classify them as strange events or as part of an open event pattern. The overview provided by the scatter plot visualization and the detailed view of values distribution in a selected set of events makes it easier for operators to execute this task quickly and without error. In addition to this conceptual analysis of visualization benefits we are developing a framework that will help us to better quantify the impact of the visualization.

5 Related Work

Others have also proposed visualization techniques for security event monitoring. Erbacher [EWF 02] proposed a method that displays event sources in a circle around event targets. Newly arrived events are animated like additional spokes of a wheel around the monitored target machines. We believe that such a visualization is more useful in settings that monitor only a few machines. In a SOC that has to cope with possibly thousands of target machines, such a visualization would require too much screen estate. We furthermore consider the generation time of an event a sufficiently central ordering property that we use it as the fixed x-axis in our scatter-plot visualization. Without this, operators must be able to animate recent changes of the visualization so that they can see recent changes in a sensor they did not monitor closely for the past few minutes.

Other visualization techniques for supporting intrusion detection have also been proposed. Girardin and Brodbeck [GB 98] proposed a visualization of firewall log entries that combines a two-dimensional display of a multidimensional clustering process with a coordinated parallel-coordinates display of the log entries. A similar visualization technique was proposed by Nyarko et al. [NCS 02]. Their visualization renders the output of a multidimensional scaling algorithm as a 3-D display, whereby clusters of similar events are formed. However, we do not think that such “self-organizing” displays of event data would be useful for SOC operators as the visualization makes it too difficult to find out quickly why some events ended up in a particular cluster. Furthermore, as new events arrive such visualizations often do not provide the kind of display stability over time that is required in practical monitoring situations. This temporal instability even makes it difficult to use animation as a tool to support operators to track the recent history of a sensor.

Finally, Hellerstein and Perng [MHP 01] presented a tool for flexible visualization of intrusion event data. Their tool also supports the kind of information displays that we provide to the SOC operators. However, the goal of their flexible visualization tool is to support the offline analysis of events to provide support for detailed analysis and knowledge discovery in the dataset. Such analysis is typically done during forensic analysis and false alarm management. Our work builds on theirs in that our basic scatter-plot and parallel-coordinates visualization could be created as a special configuration of the displays available in their system. The difference is that our system is directly connected to the real-time event stream and that the particularly chosen configuration of information displays, display options and data manipulation options is optimized to facilitate the real-time classification of events.

Other techniques to visualize multivariate data also exist (see for example [WB 97] and [KK 96]). A systematic, experimental exploration of the usefulness of these visualizations for real-time intrusion event monitoring will become feasible once we have created our experimental framework to measure the visualization impact.

6 Summary and Discussion

Operators in a centralized network security operations center (SOC) have to deal with a large number of security events collected by hundreds of sensors in the respective customer networks. The sheer amount of data makes it difficult for operators to process all the received events and monitor the stream of incoming events in real-time. We have presented a model that describes the problem-solving steps SOC-operators perform when dealing with newly arrived events. We then described an information visualization that we developed to support operators in the important initial step of a new event triage, and reported our findings from an initial evaluation with operators who use the visualization as an integrated component of a real-time console for security information management.

Besides the event monitoring console, operators in the SOC use a number of other custom-developed tools to facilitate their task and keep each other informed. For example, a specialized Lotus Notes database helps operators keep track of performed callouts. Other tools not directly linked to the primary task of event monitoring are also employed. For example, a custom Web application is used to monitor the status of network intrusion sensors.

Thus, the monitoring console and information visualization in particular is only one of the tools that can be used to support SOC operators in their problem-solving tasks. Together with other tools that support the main problem-solving process as well as additional tools for auxiliary processes, operators are enabled to manage the large number of incoming security events.

In our informal evaluation we saw that with the right set of tools, operators can be supported in mastering the problem-solving challenges in a centralized SOC. We found evidence that the current set of tools is sufficiently good that the natural benefit of centralized security monitoring can come into play. One of these benefits is the capability to perform event correlation across the customer set. Indicators of problems that occur at multiple customer sites can be spotted more easily in a multi-customer setting. Another benefit of centralized event monitoring is that information about typical attacks and attack trends can be determined with higher statistical reliability and the information can be used without delay in the online monitoring process. Furthermore there is also a natural educational benefit for centralized SOCs. Operators of centralized SOCs are more likely to be exposed to a wider range of attacks and can communicate with other members of the team, who might have gained specialized knowledge over time. Thus, we believe that with the right set of tools available, the outsourcing of security monitoring to a centralized SOC should be an interesting option to consider.

7 Acknowledgments

The text of the paper was written primarily by Markus Stolze. René Pawlitzek developed the security event monitoring console (ZEC) and Markus Stolze the visual event views. Input and ideas for improvement for the visualization came from members of the IBM Managed Security Delivery Security Operations Center as well as from members of the IBM Zurich Global Security Analysis Lab (GSAL) and the Zurich e-Business Research Group, in particular Christopher Calvert, Richard Barrera, Ben Gardelli, Stefan Hild, René Pawlitzek, Axel Tanner, Michael Walter, and Alex Wood. Andreas Wespi significantly contributed to the content and structure of the paper and the development of the event classification model. Klaus Julisch, James Riordan, Morton Swimmer and Diego Zamboni provided additional input. The ZRL publications team helped to streamline the language of the final version of this paper. All remaining language and content problems are the sole responsibility of the authors.

8 References

[EWF 02] Erbacher, R. F., K. L. Walker and D. A. Frincke (2002). “Intrusion and Misuse Detection in Large-Scale Systems.” IEEE Computer Graphics and Applications 22(1): 38-48.

[GB 98] Girardin, L. and D. Brodbeck (1998). “A Visual Approach for Monitoring Logs.” Proc. 12 th Usenix System Administration conference, Boston, Massachusetts, USA, December 6-11: 299-308. (http://citeseer.nj.nec.com/context/1508422/513384).

[ID 99] Inselberg, A. and B. Dimsdale. (1999). “Parallel coordinates: a tool for visualizing multidimensional geometry.” Proceedings of IEEE Visualization'90, Los Alamitos, CA, October 1990: 361--378.

[KK 96] Keim, D. and H.-P. Kriegel (1996). “Visualization Techniques for Mining Large Databases: A Comparison.” IEEE Trans. Knowledge and Data Engineering 8(6): 923-938.

[MHP 01] Ma, S., J. Hellerstein and C.-s. Perng (2001). “EventMiner: An integrated mining tool for Scalable Analysis of Event Data.” May 2001.

[NCS 02] Nyarko, K., T. Capers, C. Scott and K. Laeji-Osias (2002) “Network Intrusion Visualization with NIVA, an Intrusion Detection Visual Analyzer with Haptic Integration.” Proceedings 10th Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems, 24 - 25 March 2002, Orlando, Florida.

[HWG 02] P. K. Harmer, P. D. Williams, G. H. Gunsch and G. B. Lamont: (2002). “An artificial immune system architecture for computer security applications.” IEEE Transactions on Evolutionary Computation 6(3): 252-280.

[SPH 03] Stolze, M., R. Pawlitzek and S. Hild (2003). “Task Support for Network Security Monitoring.” Proceedings ACM SIGCHI 2003 Workshop on System Administrators are Users, too: Designing workspaces for managing Internet-Scale Systems. (http://www.cs.berkeley.edu/~mikechen/chi2003-sysadmin/).

[WB 97] Wong, P. C. and R. D. Bergeron (1997). 30 Years of Multidimensional Multivariate Visualization Scientific Visualization - Overviews, Methodologies and Techniques. Scientific Visualization: Overviews, Methodologies & Techniques. G. M. Nielson, H. Mueller and H. Hagen. Los Alamitos, California, IEEE Computer Society Press: 3-34.