SECURITY AND COMMUNICATION NETWORKS
Security Comm. Networks 2012; **5**:404–421
Published online 11 May 2011 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/sec.324

REVIEW ARTICLE

A survey of security visualization for computer network logs

Yanping Zhang[1], Yang Xiao[1]*, Min Chen[2], Jingyuan Zhang[1] and Hongmei Deng[3]

[1] Department of Computer Science, The University of Alabama, 101 Houser Hall, Tuscaloosa, AL 35487–0290, U.S.A.
[2] School of Computer Science and Technology, Huazhong University of Science and Technology, Wuhan, Hubei 430074, China
[3] Intelligent Automation, Inc., 15400 Calhoun Drive, Suite 400, Rockville, MD 20855, U.S.A.

ABSTRACT

Network security is an important area in computer science. Although great efforts have already been made regarding security problems, networks are still threatened by all kinds of potential attacks, which may lead to huge damage and loss. Log files are main sources for security analysis. However, log files are not user friendly. It is laborious work to obtain useful information from log files. Compared with log files, visualization systems designed for security purposes provide more perceptive and effective sources for security analysis. Most security visualization systems are based on log files. In this paper, we provide a survey on visualization designs for computer network security. In this survey, we looked into different security visual analytics, and we organized them into five categories. Copyright © 2011 John Wiley & Sons, Ltd.

KEYWORDS
network security; security visualization; logs

***Correspondence**
Prof. Yang Xiao, PhD, Department of Computer Science, The University of Alabama, 101 Houser Hall, Tuscaloosa, AL 35487-0290, U.S.A.
E-mail: yangxiao@ieee.org

1. INTRODUCTION

Because of the world's reliance on computer networks, their security and stability are of tremendous importance. A computer network is vulnerable to attacks and errors, and many attacks, such as a worm attack, a denial-of-service attack, and so forth, can lead to billions of dollars in damages [1]. Even if a particular network system has strong enough protocols, flaws and instability may still happen during deployment [2]. Two kinds of threats are differentiated here. Flaws refer to unintentional errors in the system, whereas intrusions refer to intentional illegal accesses or even malicious attacks on the system [2]. Therefore, there is no perfectly secured system, and intrusion/flaw detection is vital for the protection of computer systems [2].

For the security and the stability of computer network systems, log files are the main resources for analyzing the performance of the computer network [2]. System logs are critical for administrators to obtain the security information of the network. Generating and continuously updating log files are tracking mechanisms for operating systems or software to record all of the activities of the running system [3-76, 134-140]. For example, in Windows operating system, log files about security, system, application, and so forth, are available [3].

However, it is a difficult task to search for problems through analyzing logs [2]. Log files are not user friendly, and not all of the information is useful for detecting abnormal activities [3,7]. Meanwhile, some applications generate multiple log files to record categorized data. Therefore, the analysis must combine several log files [3].

As a branch of computer science, information visualization is such a research field with the assistance of some professionally designed software tools to model complex data with interactive images [77]. Administrators can easily learn, understand, and recognize threats, vulnerabilities, and attacks, which are the basis for further response [77]. Many researchers have provided their visualization design for security purposes, and many software tools are designed for monitoring purposes.

In this paper, we study different security visualization designs and provide a categorization based on the forms of visual views.

The rest of this paper is recognized as follows: In Section 2, we introduce the data process and analysis for security. In Section 3, we provide the classification of visualization designs, and, in Section 4, we study different visualization designs. In Section 5, we discuss some future research directions. Finally, we conclude the paper in Section 6.

2. DATA PROCESS AND ANALYSIS

In order to maintain the security and the stability of computer and network systems, vast amounts of data are needed for dynamic analysis. Data are the main resources for security and forensic analyses. Regardless of what schemes an analysis system employs, log files are the basis for analysis. As described in the former section, it is a laborious task to obtain useful information for anomaly detection. In this section, we introduce and study two methods of data processing and analysis: data mining and visualization analysis.

Data mining refers to extracting patterns from data [78]. As amounts of data are doubling very quickly, data mining becomes an increasingly important and effective tool for transforming these data into information. It is widely applied in various fields, such as business, marketing, monitoring, attack/intrusion detection, and knowledge discovery in the real world [78].

"A picture is worth a thousand of words." Information visualization generally needs to handle very large amounts of textual, symbolic, or relational data and to transfer these data into graphics that can be displayed [79]. A visualization system provides a more perceptive method for security analysis. The advantage of visual analysis is that it is easier for humans to find unexpected patterns through pictures.

2.1. Data mining

Data mining provides an effective method for data sifting, which can help filter out data with certain patterns. Attacks/intrusions with regular patterns can be easily detected through data mining technology [80]. Data mining technology is widely applied in marketing, finance, fraud detection, telecommunications, and data cleaning [80].

When traditional signature-based intrusion detection methods became ineffective in discovering novel attacks, data mining methods began to play a significant role in intrusion detection. Data Mining is defined as the extraction of patterns or models from data sets [3] and known as Knowledge Discovery in Databases (KDD) [81,82]. Different data mining methods are employed in different approaches in which different data are analyzed. For example, Schultz *et al.* [83] utilize Naive Bayes algorithms for the detection of malicious activities, while Ghosh and Schwartzbard [84] apply neural networks on the US Defense Advanced Research Projects Agency (DARPA) network connections dataset.

With specific algorithms, data mining technology can extract patterns from data [80]. Some additional steps in the KDD process provide more accurate results, such as data preparation, data selection, data cleaning, incorporation of appropriate prior knowledge, and proper interpretation of the results of mining [80]. These steps are essential in ensuring that the derived information from the data is useful and correct knowledge [80]. KDD is an interactive and iterative process involving many steps that require decisions by a user [80]. Some data mining methods blindly process data and easily lead to meaningless and invalid patterns, and this is a dangerous activity [80].

With the extensive application of data mining technology to uncover patterns, an important consideration is the data sample. If the data samples are not representative enough, the sample may produce indicative results of the domain [78]. Meanwhile, if the "mined" data sample doesn't have the pattern that we are interested in, it might be assumed that the pattern is not included in the whole data. The selection of the data sample is critical for the data mining [78]. Similar to any other analysis, data mining only functions in conjunction with the appropriate raw material [78]. Users must first collect indicative and representative data. Furthermore, a specific pattern discovered in a sample of data does not necessarily mean that the pattern is representative of the whole population that the data set was drawn from [78]. Therefore, verification and validation of the pattern is an important part of the data mining process [78].

In [85], an algorithm to extract patterns is presented, and a method is designed for comparison with those patterns, through which mismatches could be found and alarms could be raised once abnormal activities are detected. In [86], Mahoney and Chan employ another data mining technique, association rules, for intrusion detection. Data mining techniques are also helpful for computer security. An example is that in a work by Michael [87] utilizing suffix trees to get frequent series of system calls.

Similar to data mining, many methods are used in intrusion detection. In [88], two domain-independent online anomaly detection schemes are designed; namely, the Lempel–Ziv (LZ)-based and Markov-based detection schemes were proposed by exploiting the location history traversed by mobile users. In [89], a Markov chain-based approach and a Hotelling's T2 test-based approach were proposed. In [90], by exploiting regularities demonstrated in users' behaviors terms of calling and mobility activities, the authors presented a suite of detection techniques to identify fraudulent usage of mobile telecommunication services: a nonparametric technique known as the Parzen window with a Gaussian kernel, which is used to estimate a class-conditional probability density function, and a Bayesian decision rule, which is applied in order to achieve a desirable error rate.

2.2. Visualization

Visualization analysis is an interactive process, a continuous loop between visualization and knowledge discovery, and sometimes even goes back to data collection/preparation [2]. At first, a visualization system is developed based on analysis goals and collected datasets.

With the exploration of the dataset, users get some previously unknown knowledge related to the data. Further questions and analysis may then be required using current tools or after acquiring new tools. Such a loop ends when the user is finally satisfied with the obtained information. A visualization system helps users detect patterns in a more perceptive way. When combined with more technologies, such as data mining or machine learning, a visualization system can more effectively achieve patterns.

The exploration of the data visualization system helps us explore and analyze data under study. It is an inherently iterative process that includes multiple steps [79]. As a result of explosive development of Internet technology in recent years, information visualization [91] has become a significant research area of computer science. However, the application of visualization in computer security is still few. These include Erbacher *et al.* [92] employing glyphs in visual intrusion detection data, Yurcik *et al.* [93] designing a tool for the visualization of the network traffic, Girardin [94] working on packet-based visualization, and Tudumi [95] designing a visualization system for the surveillance and audience of computer logs in order to assist in the detection of malicious activities.

Data mining methods for visualization purposes are also few in number. Perception-based classification (PBC) [96] is one of them, which is particularly related to computer security. As a visual tool for classification, PBC can be applied to detect intrusions. Meanwhile, PBC [96] is an improvement of data mining methods by the successful combination of visualization and machine-learning techniques. Another example is MineSet [97], which is an integrated analytical and visual data mining tool. It provides an interactive exploration process to find trends and relationships among data [97]. Mielog [98] is a forensic analysis tool specifically for system log files. It statistically analyzes and classifies log entries. This visualization system was designed to demonstrate the features of logs but not the contents.

There also exist some other network visualization tools, such as the Internet Mapping Project [99] and the H3Viewer [100]. Munzner *et al.* [101] worked on the visualization of the global topology of the MBone. These tools focus on the demonstration of changes in reachability and topology. Labovitz *et al.* [102] also conducted some simple visualizations of the number of Internet routing changes.

In [103], Goldstein *et al.* described an iterative and interactive data exploration process. It is an interactive anomaly detection system.

In [104], a geolocation system is introduced. The function is realized based on a software system named StoneGate Management Center. The system tracks where packets come and where they go based on a global IP address database and the global location map. It can statistically calculate the top rate of all hosts and locate a specific IP address. The corresponding communication of hosts is also recorded in the log file and can be demonstrated when users need it.

In [105], a hierarchical visualization method based on IP address is designed. This is a visualization method that hierarchically demonstrates each IP address and its activities.

Various ideas are employed to design visualization systems. The purposes are the same: to help administrators detect malicious activities and anomalies. In Section 4, we will introduce different designs in detail.

3. CLASSIFICATION OF VISUALIZATION DESIGN

In [106], visualizations for security analysis are categorized into hierarchical visualization techniques and nonhierarchical visualization techniques. In [107], authors classified the visualization methods for wired networks based on presentation mode, as shown in Table I. Different presentation modes include text based, dashboards, chart and graphs, and visualization. Currently, visualization is the most popular presentation mode.

Security visualization for wireless networks is focused on in [107]. Current studies are mostly based on the wireless network mapping technology, which intends to show the location of Access Points (APs) and mobile devices as well as their links [107]. There are two main wireless network mapping methods: the GPS and radio frequency (RF) signal-based methods [108].

The GPS-based method uses the conveniences of GPS, which can be utilized all around the world [107]. However, their surroundings may greatly limit the performance of GPS receivers. For example, having tall buildings around, being indoors, or being under cover may make it difficult to get position information because of limited satellite signals [107].

Radio frequency signal-based methods utilize varieties of RF signals like wireless LAN (802.11), GSM, and so forth. Most approaches use the received signal strength to do the locating work [107].

In this paper, we categorize all kinds of designs based on the forms of visual results. We do not separately study the visualization technology for wired or wireless networks. We classify visualization designs into the following five categories: text-based visualization, parallel visualization, hierarchical visualization, three-dimensional (3D) visualization, and other designs. We will introduce each category in Section 4.

Table I. Classification of visualization of wired network [133].

Presentation mode			
Text based	Dashboards	Charts and graphs	Visualization

4. VISUALIZATION DESIGNS

We classify visualization designs based on the form of visual results. In this section, we study different visualization designs in each category.

4.1. Text-based analysis

The text-based method is the most traditional and fundamental method for security and forensic analyses. Many tools designed to analyze log files use a text-based view. We will study some traditional tools, such as an MS Log Parser (Microsoft Corp., Redmond, WA, USA), as well as some novel ideas like geolocation, which associates log data with global location information.

4.1.1. Geolocation.

The StoneGate Management Center [104] was developed to monitor activities of specific IP addresses, as well as to locate the geographical locations of IP addresses in real time. A geolocation map, as shown in Figure 1, is generated based on the public geolocation database, which is built and updated in an internal management server [104]. The software cannot only trace the locations of IP addresses, but can also track where packets are from and where they are going.

There are also options to show the top rate statistics as a map. Geolocation can show the top rate statistics in a pie chart view as well as a geographical view. It can also show the log records of a host through right clicking on the IP addresses.

The geolocation map is also conjunct with Google maps. With the Google maps view, it is easier to find the specific location of a host. Geolocations can also be configured for private IP addresses. The only difference is that internal IP addresses need to be configured manually. Then, all of the traffic within the network can be monitored.

4.1.2. Log Parser.

The MS Log Parser is a powerful tool designed to access text-based data like log files, XML files, and CSV files [109]. Structured query language statements can be used with this tool, along with more powerful functions like sorting, filtering, gathering data from an input format, and achieving results of target format for visualization [3,109]. Table II shows an example of a report generated by Log Parser [3].

4.1.3. Some wireless network tools.

For wireless network visualization, most methods used text-based or dashboard visualization, whereas only a small amount of researches concerned visualization design [107]. Examples of tools using text, dashboard, or chart visualization are Netstumbler [110], Commview for Wifi [111], and WirelessMon [112], respectively. By scanning the communication channels, these tools capture wireless packets. After analyzing these packets, related statistical information is obtained. The tools mainly collect the information of APs, such as security set identifier and signal strengths.

The text-based system has great limitations, especially when there is a large amount of information. That is why visualization approaches are developed. Examples of tools using visualization as a presentation mode are IntraVue [113], Wi-Viz [114], and WVis [115].

Table II. An example report generated by Log Parser [3].

Computer name	Time	Username
PC1	8/15/2007 9:25	Administrator
PC2	8/15/2007 10:37	Local service
PC3	8/15/2007 11:10	Jacica
PC4	8/15/2007 0:25	Jone
PC5	8/15/2007 13:48	Rose
…	…	…

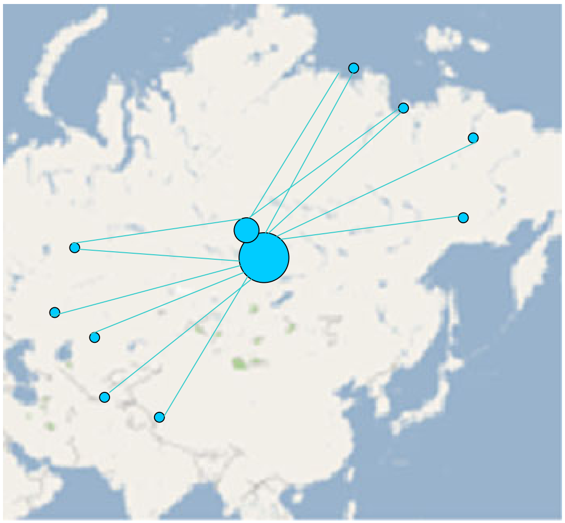


Figure 1. Geolocation map [104].

DOI: 10.1002/sec

4.2. Parallel visualization methods

4.2.1. Picviz.

Picviz [116] is a tool designed to improve the output image based on acquired data, such as logs. It employs parallel coordinate technology. Parallel coordinates [117] are a way to illustrate an event carrying *N* kinds of information within a two-dimensional (2D) plane (commonly *N* is larger than 4). Although *N*'s dimensional vector is difficult to plot, this method provides a neat and easy solution [116].

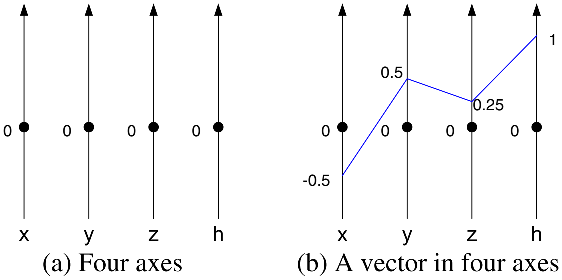
For example, a four-dimensional vector can be drawn in a 2D plain as in Figure 2(a). An example vector (–0.5, 0.5, 0.25, 1) is shown in Figure 2(b).

Superficially, a complicated structured pattern is introduced and many points correspond to many polygonal lines overlapping each other [107]. However, there is a certain relationship between the points and such a pattern. Figure 3 shows the relationship between a line in a traditional 2D (*x*,*y*) plane and a parallel plot system.

Log files are the most important data sources in this paper. Through the data acquisition process, logs are transformed into a PCV file, which is a format that can be handled by Picviz to do the visualization work. System logs, application logs, databases with information structurally stored, and so forth, can be used for data acquisition.

Figure 4 is an illustration that is similar but not identical to the visual results provided by Picviz. It is the visualization of an example of the auth syslog facility.

The first axis is the time with 00:00 at the bottom and 23:59 on the top [118]. The blank area in the first axis shows that there was no access during that time period. The exact time period in this example is from 2:29 AM to 5:50 AM. The second axis shows the machine where the log file is from. Lines go to a single point in the machine (second) axis, and this means that all of the activities are recorded in the same machine.



(a) Four axes (b) A vector in four axes

Figure 2. Two-dimensional plain for *N* dimensions [116].

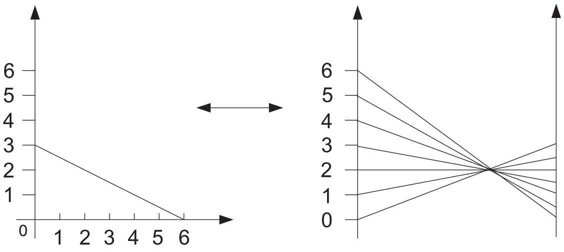


Figure 3. Two-dimensional line relationship [116].

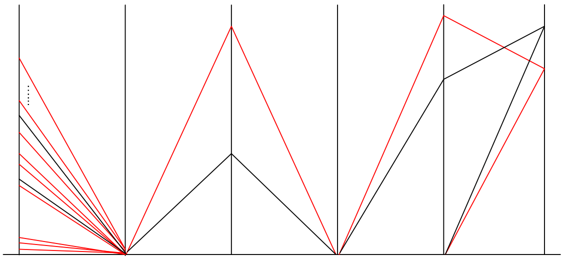


Figure 4. Picviz front end showing pam sessions opening [118].

The third axis shows what exact service or application wrote the log item. Taking this figure as an example, when the mouse is placed on the red line at the service on the top, the "su" service will be shown, which is used to log in as a root.

The fourth axis shows what pam module was used to perform the login authentication. In this example, there was only local authentication using the pam unix module, so the lines converge to a single point again. Other kinds of authentication may also occur, such as remote authentication.

The rightmost two axes show the user source and the destination of the logs. This visualization system can help administrators find abnormal activities and which IP addresses are involved.

4.2.2. Visual firewall.

Firewall [119] provides a real-time traffic view for users to certify configurations of their firewall and to monitor their network activities passively.

The visualization uses a perceptually parallel process technique. Four totally different views are established, and we only show the real-time traffic view and visual signature view as an example [119] in Figure 5.

Figure 5 [119] is an illustration of the real-time traffic view. Both incoming and outgoing packets are shown by glyphs. The direction of the traffic is demonstrated by motion, which also illustrates whether or not the traffic was rejected [119]. The packets flowing between the firewall and foreign IP addresses are demonstrated in this view [119]. In this figure, firewall rules can be verified because both accepted and rejected packets are easily to be read from the figure [119]. The left axis always shows the port number of the local host [119], whereas the right axis denotes the foreign host IP address and port number [119].

As shown in Figure 6, packet flows are plotted as lines between two parallel axes. The left axis shows ports on the local machine, and the right axis shows the global IP address. When the local host connects to a foreign host, a line connecting the local port to the global IP address will be drawn between the two axes to track packets [119].

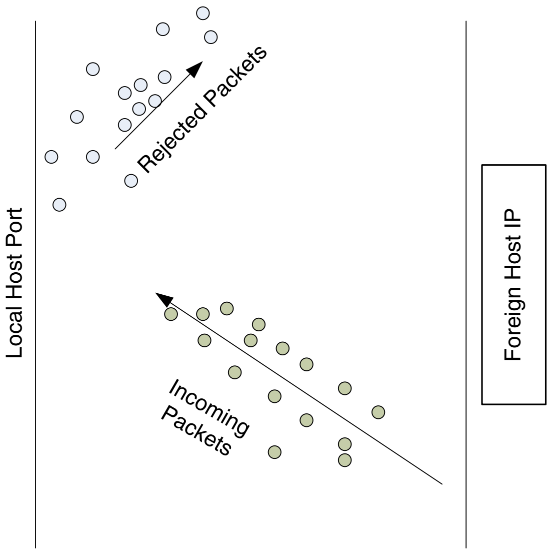


Figure 5. Real-time traffic view [119].

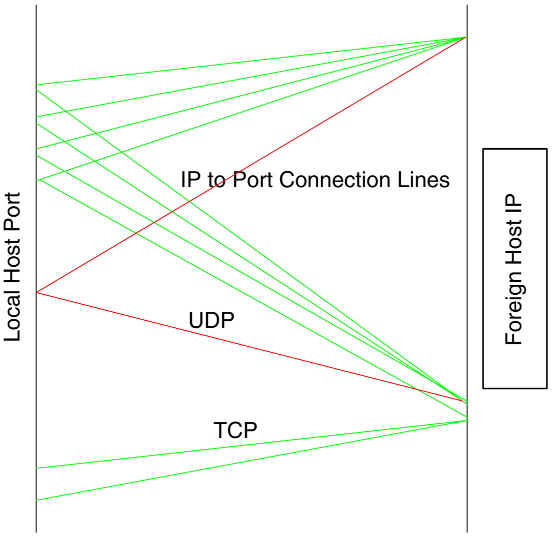


Figure 6. Visual signature view [119]. TCP, transmission control protocol; UDP, user diagram protocol.

Two colors are used to distinguish transmission control protocol (TCP) packets (green) and user diagram protocol (UDP) packets (orange) [119]. Lines fade over time. A newer line is brighter. This allows time to be demonstrated in the view [119]. This view provides a clear view for both incoming port scans and outgoing ping sweeps.

4.2.3. Rumint.

Rumint [120] is a parallel visualization tool. A data set captured over 5-min post-infection of a sandbox victim with a typical Storm variant is visualized in Figure 7 [120].

As shown in Figure 7, an illustration of the visualization results based on the data in [120], the number of axes is 6. Users are able to choose as many as 19. The six parameters shown in the figure are ideal for visualizing Storm. These parameters are the following: Packet Length, Source IP, Dest IP, UDP Source Port, UDP Dest Port, and TTL [120].

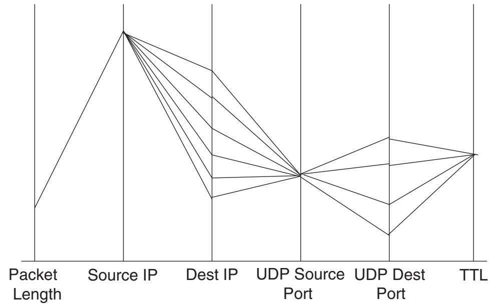


Figure 7. Rumint's exquisite detail [120]. TCP, transmission control protocol; TTL, Time To Live; UDP, user diagram protocol.

4.3. Hierarchical visualization method

4.3.1. Treemaps.

Treemapping [121] is a technology to hierarchically (tree-structurally) display data using a set of nested rectangles. Each branch of the tree is associated with a rectangle and smaller rectangles inside represent sub-branches. Treemaps [122] can be used for many different analysis purposes. Figure 8 is an illustration, which shows an example of firewall log analysis with treemaps.

For each source address, a gray block is established to demonstrate all of the destinations that are connected. For example, the IP address "195.141.69.42" is connected to the machine with IP address "239.255.255.253". Meanwhile, for each destination, another block is set up to demonstrate all activities that the source machine tried to access. Different colors indicate whether the connection is blocked (red) or not (green) [122]. Then a nice view of the firewall activity is demonstrated, which can help us detect attacks, wrong configurations, and so forth. From Figure 8, it is easy to see that there might be something wrong with the configuration of "212.251.86.126". It tried to connect a

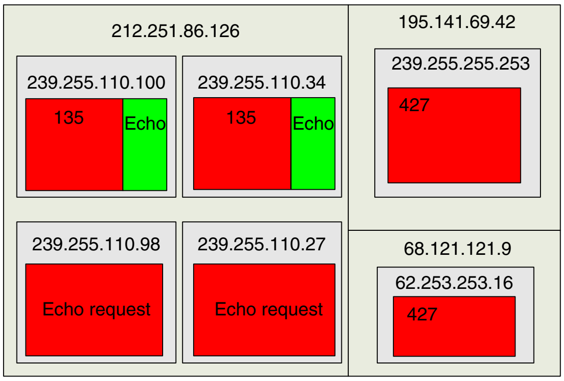


Figure 8. Firewall log analysis using Treemaps [122].

lot of machines through port 135. If this failed, it would not try again. The machine needs to be reconfigured.

4.3.2. A hierarchical visualization method based on IP address.

A hierarchical visualization method based on IP addresses is introduced in [105]. As shown in Figure 9, each IP address is displayed as a black square icon, and the whole network is displayed in a hierarchical style. As shown in Figure 10, each icon has a height that shows how many incidents are related to the IP address. Different colors of the histograms distinguish sent incidents and received incidents.

4.3.2.1. Hierarchical structure. In the hierarchical process, computers are grouped by their IP addresses, starting with the first byte of the IP address, then the second, and finally the fourth byte [105]. All IP addresses are grouped into a four-level hierarchy. The grouping process is shown in Figure 11, where the black icons show a computer and the border lines represent groups [105]. The system can visualize thousands of icons in one display space without overlapping. Clearly, zoom functions are provided for users to access details [105].

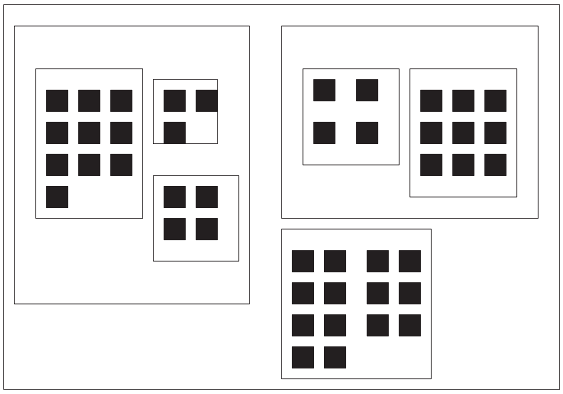


Figure 9. Example of hierarchical visualization [105].

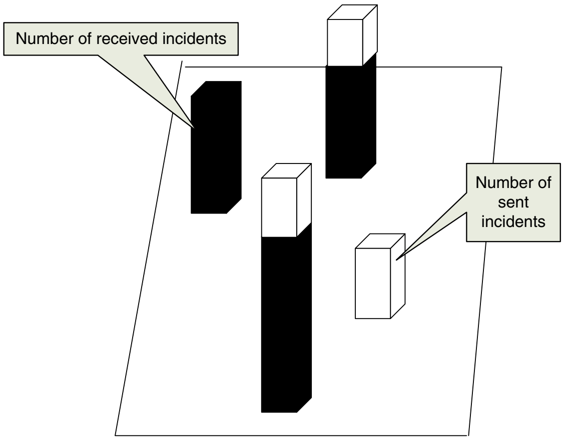


Figure 10. Each leaf node heights denotes the number of incidents [105].

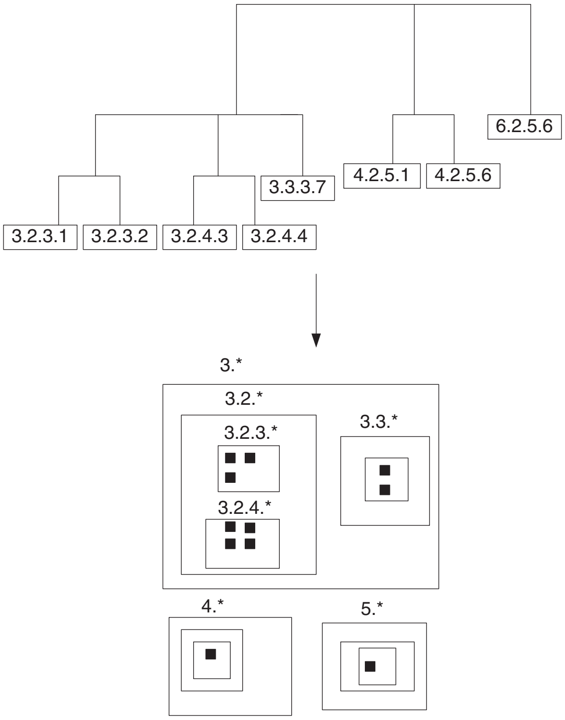


Figure 11. Hierarchy of computers based on IP addresses [105].

Usually, IP addresses are assigned based on the organization's structure. Therefore, the hierarchy reflects real computer's organization of different departments. Therefore, the technology can show the relationships between incidents and organizations [105].

4.3.2.2. Visualization analysis. In [105], experiments are conducted using the log files of Cisco Secure IDS 4320 (Cisco Systems, San Jose, CA, USA), a commercial intrusion detection system (IDS). Graphical user interface is developed using Java (Sun Microsystems, Santa Clara, CA, USA) for users to do configuration. Based on the data organization of the log files, users can choose a specific date, time range, security level, IP address, and so forth [105].

In an example in [105], a sequence of visualizations is demonstrated based on the log file recorded over 6 h including about 4000 computers to show multiple senders and receivers lasting 5 min (it might be the case that some attackers randomly select some machines as targets), concentrated incidents from a single sender to a single receiver (a sender focuses on attacking a specific computer), the former sender being disconnected and new senders arising (although the sender was blocked, several new senders arise and attack new receivers, and one new receiver locates in a different department from the continuously receiver), and many receivers in the same department being attacked (some former senders and receivers have been blocked, whereas many new receivers in the same department are attacked in a short time). Such an attack is considered to be a scan attack.

4.4. Three-dimensional visualization

4.4.1. InetVis.

InetVis [120] is a 3D scatter-plot visualization tool for network traffic. In one example in [120], the setting interface of InetVis's control panel and visualization results are shown, where the results are from the data set that was captured over 5-min post-infection of a sandbox victim with a typical Storm variant: the horizontal blue x-axis denotes the destination address (home network), the red z-axis denotes the source address (external Internet range), and the green y-axis demonstrates ports (TCP and UDP) plotted along. In the example in [120], the infected host is a single point of reference in the red z-axis. The IP address of the host is "192.168.248.105", which is a Class C address [120]. Hundreds of hosts were infected, which is demonstrated through the prism of visualized disease across the blue x-axis [120]. Meanwhile, the rainbow across the y-axis (green) shows the port range.

4.4.2. An integrated visualization system.

With the increase of all kinds of attacks, like worms, bonet, and so forth, IDS is becoming more and more important. However, a serious drawback of current IDS logs is many false positives [123], which are alerts for illegal activities. As visual patterns are already set up for typical attacks, it is easier to use a visualization system to detect them than text-based methods [123].

Usually, attacks refer to illegal activities from an external network [123]. When detecting worms, administrators can take appropriate actions to avoid these attacks. However, internal attacks on external networks are also serious problems. Some machines infected by a virus may continuously attack other machines within the internal network as well as machines in the external network [123]. For internal attackers, administrators are always interested in the locations of the infected machines because these machines might need to be disconnected or manually powered off [123].

An integrated visualization system for a large-scale local area network is described in [123], which can be used for external attack detection as well as internal monitoring.

4.4.2.1. Visualization with three planes.

Three planes are displayed in the visualization system totally. As demonstrated in Figure 12, the planes are time graph, IP Matrix, and map.

In the time graph plane, the horizontal axis is the time axis, which increases on the right side, and the vertical axis is the amount of detected attacks [123]. Different colors are used to distinguish different lines. Meanwhile, the lines of the time graph are able to move along the plane of the IP Matrix because they have different IP addresses [123]. Usually, the exact value of the attack is not as important as the changes or similarities of the lines [123].

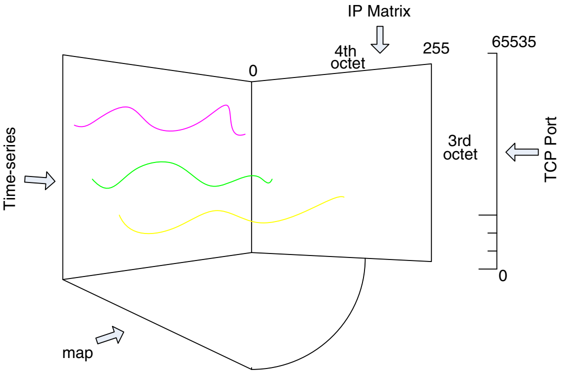


Figure 12. Example overview of the system [123].

The IP Matrix shows two octets in its horizontal and vertical axes. If the first and second octets are in the IP Matrix, the system is functioning to detect external attacks. When the third and fourth octets are demonstrated, the system is monitoring the personal computers (PCs) within the local area network [123].

Finally, the third plane is the geography map of the local network. This plane is only useful for the administrator to locate PCs in the local network when malicious activities are detected within the local network. Therefore, this plane is only useful for internal network monitoring. The map is configured to be connected to IP addresses [123]. When the system needs to locate a specific IP address, a line will be drawn from the IP Matrix plane to the map plane [123].

An additional axis is displayed on the right side of the graphical user interface, which is the TCP port. Usually, the port number tells us what service is attacked [123].

4.4.2.2. Two examples: botnet and Secure Shell brute force attack.

Figure 13 is an illustration that shows an example in which the system detects alerts on Internet relay chats, a well-known communication channel

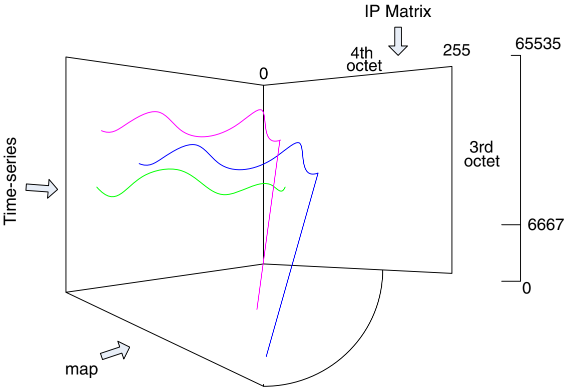


Figure 13. Visualization of botnet activity [123].

of botnets [123]. From the figure, we can learn that port 6667 is attacked, which is often used by the Internet relay chat servers [123]. The blue and purple lines show similar results. Both of the lines show the result of botnet and that the two infected machines are located in different departments, as shown in Figure 13 [123].

Figure 14 is also an illustration that shows an example of attacks to TCP 22 port (SSH port) [123]. Secure Shell (SSH) brute force attackers repeatedly connect to and try to log into another computer [123]. As there are very many connections, the result figure is somewhat unclear.

4.4.3. Flamingo.

Flamingo (Merit Network Inc., Ann Arbor, Michigan, USA) [124,125] is a tool that can visualize Internet traffic data in real time. Figure 15 is an illustration that shows an example of a sequence of packets that are sent from a subnet to a variety of IP address. The left side denotes the source IP address, and the right side denotes the destination address [125]. The figure seems to be fan out from left to right, indicating a network wide scan [125].

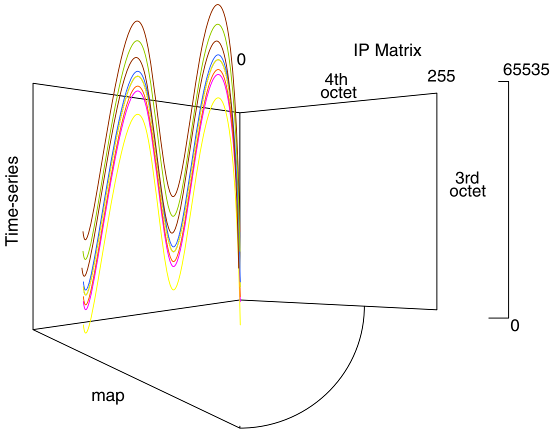


Figure 14. Visualization of attacks to SSH [123].

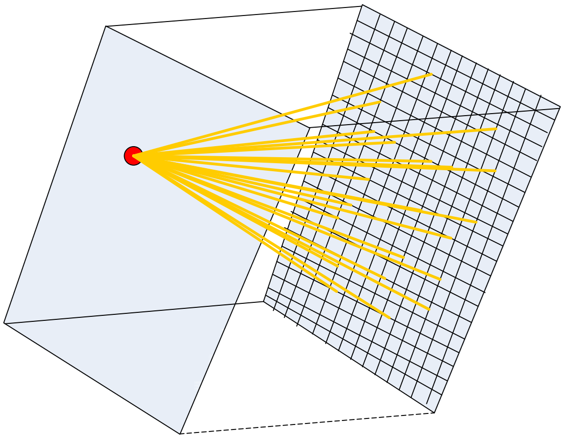


Figure 15. A scan attack from a subnet [125].

4.4.4. MineSet.

MineSet [97] is an integrated analytical and visual data mining tool. It provides an interactive exploration process to find the trends and the relationships among data [97]. Figure 16 is an illustration of results of experiments in [3]. A bad status code from the access log file of a web server is detected in the experiment, which includes the records of dozens of days' continuous operation [3]. The highest black bar demonstrates 101 bad status of a client with a specific IP address. The highest gray bar demonstrates another client with 50 bad status codes [3]. Through the visualization results, administrators are able to find clients with malicious purposes and block their requests in time [3].

4.4.5. A hybrid intrusion detection and visualization system.

A hybrid intrusion detection and visualization system is introduced in [126]. According to the classification in [126], there are three main detection categories: signature detection, anomaly detection, and hybrid detection. Signature detection works based on the signature matching of known attacks [126]. As an example, worm attacks can be detected by signature detection because of the over use of network services and resources. However, it is difficult to detect novel attacks through signature-based detection [88–90,126,127]. Anomaly detection works with assistance of models of normal behaviors to identify anomalies in computer system performance metrics, such as I/O activity and CPU usage. Hybrid detection is a method that combines the advantages of both methods.

The authors in [126] tried to establish a systematic and intelligent visualization system that includes a two-stage intrusion detection technique. The first stage employs a signature-based detection method to detect intrusions. A database of known intrusion behavior is established and updated over time. Data mining is quite effective for this stage's work. During this stage, audit data are compared with the database in real time [126]. Once an intrusion is detected, interventions and precautions will be taken based on different mechanisms. Meanwhile, the system call information will be shown in a graph for further analysis [126].

The second stage, the anomaly detection stage, aims to detect novel attacks [126]. Additional detection is provided, such as the access of confidential data. When monitoring a specific program, the system will compare the event traces with some expected behaviors [126]. The authors characterize a program through a series of system calls [126]. A safe range of system calls is established in the system, and a fuzzy inference approach is employed. In one example in [126], a normal value of system calls is shown with the relationship of system calls, the number of system calls, and time. The fuzzy inference module works under a sequence of rules. The rules are based on prior knowledge of the system. For example, a high value of Syscall2 and a low value of Syscall3 are abnormal. The prior knowledge can be established through a learning

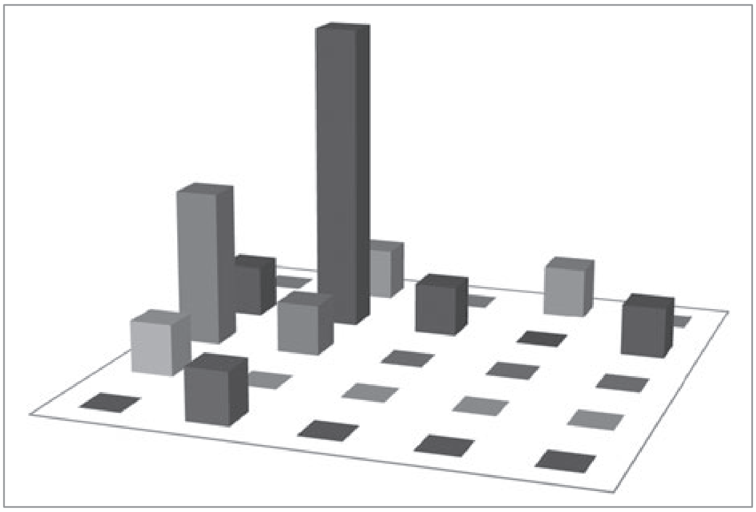


Figure 16. Total count of bad status codes for each IP client [3].

method such as neural network [126]. When an anomaly is detected, it will be reported to administrators for further actions [126].

4.5. Other designs

4.5.1. A visualization system based on monitoring border gateway protocol data.

Anomaly detection refers to the detection of malicious behaviors [128]. Mostly, an anomaly in network data is recognized by statistical analysis, where statistical measures are used to model normal profiles [129,130]. In [128], the authors designed a visualization system, in which the anomaly is detected quickly without establishing normal data set.

4.5.1.1. Border gateway protocol. Border gateway protocol (BGP) [131] is critical for communication among routers in order to maintain network connectivity. Therefore, analyzing BGP data is a good way to understand the behaviors and performance of the Internet, which also helps users determine the characteristics of specific routing activity, the weaknesses in the network, and even anomalous behaviors [128].

4.5.1.2. As routes and origin autonomous system changes. An IP prefix is used to identify a network in the Internet [1]. As an example, IP prefix “202.210.0.0/21” demonstrates a network within which each IP address has the same first 21 digits. An autonomous system (AS) refers to a single administrative domain including one or more networks. Each AS has a unique AS number. Usually, the cluster of machines included in an AS shares portions of their IP addresses.

Border gateway protocol is the protocol used to exchange reachability information between two ASs, based on which packets can be forwarded to the correct destination. BGP routers communicate through a form of BGP announcement. An example of BGP announcement is “202.210/16: (6,14,27)”, which means that, by going to AS-6, AS-14, and finally AS-27, the IP prefix “202.210/16” could be reached.

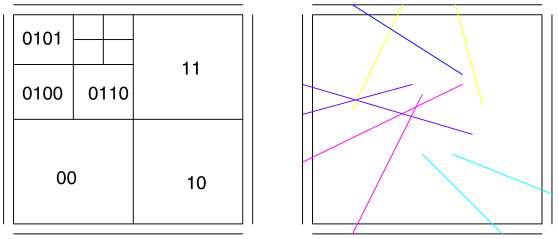
The last AS in the path to reach an IP prefix is the origin AS of the prefix [1]. In the former example, the origin AS of the IP prefix “202.210/16” is AS-27. Different reasons may lead to a change of the origin AS, such as the ownership change of the prefix, misconfiguration of routers, or malicious attacks.

According to daily BGP data, the origin AS change (OASC) can be recorded. In [128], OASCs are recorded as entries with the following format (Prefix, AS, Data, Type).

Origin AS changes are mainly classified into four types, which are further categorized into eight types [79]. The four main classes are: B-type, H-type, C-type, and O-type.

C-type refers to an AS announcing a prefix previously owned by another AS [79]. C-type is further classified into the following four types [79]: CSM (a C-type change from a single AS to multiple origin ASs), CMS (a C-type change from multiple ASs to a single origin AS), CSS, and CMM. Some OASCs are complementary. For example, a CMS event could correct a CSM event.

4.5.1.3. The visualization system design. In [128], an OASC visualization system has been created. The mapping scheme for data values to be mapped to graphical values is as follows: IP prefixes are mapped to pixels on a square. As shown in Figure 17 [128], the mapping is done in a quad-tree manner. Each square is continuously divided into four equal squares [128]. For example, to demonstrate a 32-bit range address within a



(a) Mapping of IP prefix (b) OASCs with different color

Figure 17. The plain design for visualization [128].

square, the mapping begins at the first two most significant bits of the 32 bits, through which it is placed into, of the four sub-squares [128]. The prefix is then put into a sub-square within a sub-square. The division is repeated until all of the bits of the prefix are exhausted [128]. In [128], a 512 × 512 pixel square is used to represent the entire IP prefix space. As IP prefixes have at most 24 digits as masks, at most 64 different IP prefixes may be represented by the same pixel.

Meanwhile, four lines are drawn surrounding the IP square in order to demonstrate AS numbers in the network. All of the AS numbers are mapped to a pixel on one of the four lines, as shown in Figure 17. If an OASC occurs, a corresponding line will be drawn connecting an IP prefix with an AS number. Each pixel on the four lines represents more than one AS number [128]. Zooming features are provided in the main display for users to distinguish different AS numbers. Meanwhile, the colors of the lines are distinguished according to the type of OASCs [128]. For example, yellow line denotes CMS, and a green line denotes CSM.

Consecutive data of a day can be shown as a "movie" or frame by frame, which can help the user detect temporal patterns [128]. Meanwhile, a certain amount of the previous days' data can also be shown on the window to assist the user's memory of patterns from previous days [128]. The displays of previous days data are darker, and users can control which days are shown [128].

Figure 18 is an illustration that shows an example of CSM events and corresponding CMS events to correct the abnormal status [128]. The first figure shows that, on 6 April 2001, a large number of CSM events exist because of a specific AS announcing prefixes of many different ASs. After that, a sequence of CMS events happened to correct the errors, and the system finally went back to normal on 13 April 2001.

4.5.2. A visualization design for monitoring e-mails.

In [132], a visualization system for monitoring e-mails containing inappropriate content is designed. This visualization is designed for supporting human resources (HR) to manage e-mails with inappropriate materials.

In the visualization technique, all internal senders of inappropriate materials are mapped around a circle. Then, a line with an arrow will be drawn cross the circle area from the sender to the receiver. Senders are differentiated by the color of the lines. If the monitored e-mail is from an external sender, a small arrow line is used outside the circle looking like the recipient is sending the e-mail. This will help us to determine whether the employee forwards inappropriate contents after receiving them from the outside. By tracking the e-mails in this way, a chain of e-mail forwarding can be made to find who has forwarded them.

The designer manually maps related data with Microsoft Visio (Microsoft Corp., Redmond, WA, USA) and uses the layering capability to see each sender's own Visio layer. Figure 19 shows an example of tracking a specific e-mail, and Figure 20 is an example of the activity of a specific sender.

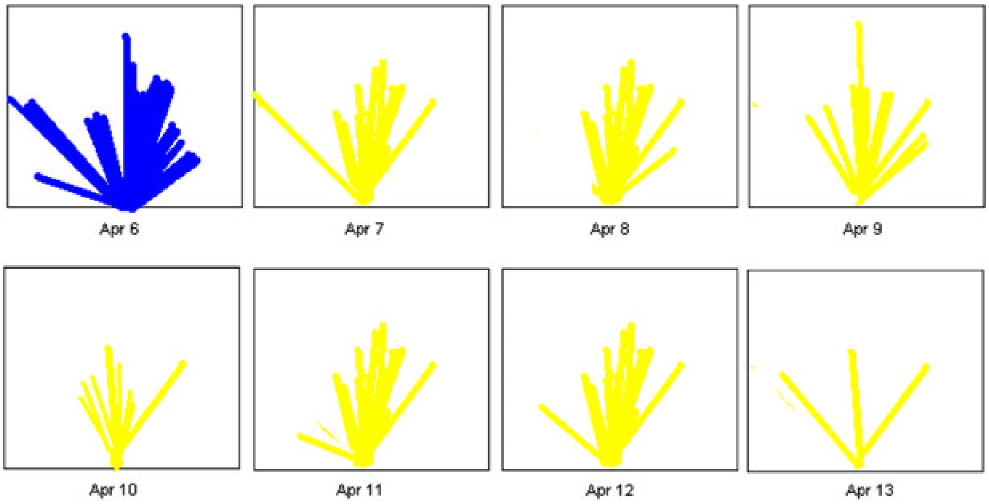


Figure 18. CSM activity on 6 April 2001 followed by 6 days of corrective CMS activity [128].

DOI: 10.1002/sec

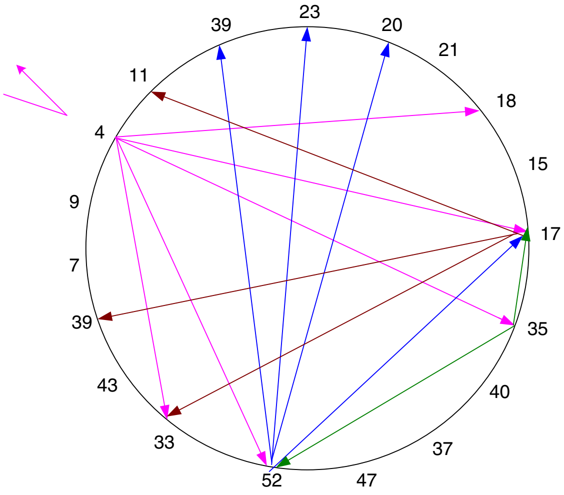


Figure 19. The tracking of a specific e-mail [132].

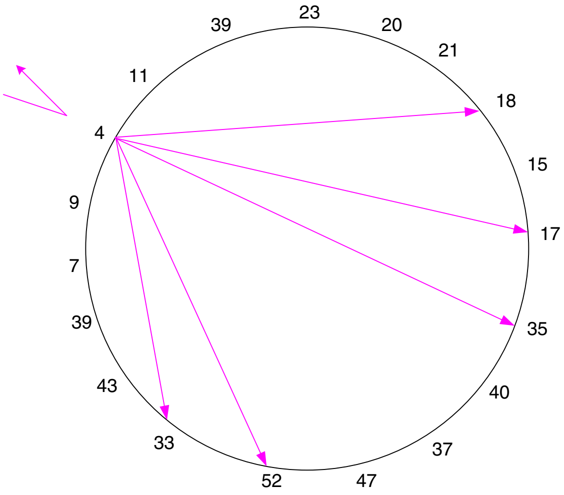


Figure 20. The activity of a specific sender [132].

This visualization system cannot only detect how many people are involved in forwarding e-mails with inappropriate contents but can also calculate how many inappropriate e-mails are forwarded by a specific sender.

5. FUTURE RESEARCH DIRECTIONS

An important characteristic of visualization design is that there is no fixed model to transfer datasets to spatial coordinates. Therefore, there is no restriction for mapping data attributes to visual attributes. As we have studied, all kinds of mapping methods can be utilized. Parallel coordinates are utilized in [116] to illustrate *N* kinds of information within a 2D plane. In [105], hosts are mapped to a hierarchical visualization according to their IP addresses. The principle is to make discoveries through appropriate mapping.

Our future work is to design a visualization system for network monitoring. Our mapping technology will combine the hierarchical visualization with the parallel coordinates. The hierarchical technology [105,122] provides a clear view of each host and corresponding results of incidents. However, the visual results are limited, and many attributes are not clear to users. Parallel technology [116,119,120] provides an easier way to show multiple attributes, but the representation of axes may lead to messy visual results. Therefore, we will extend the forms of visual results. Multi-dimensional visualizations such as 3D views and parallel coordinates will be provided to depict patterns or correlations among attributes.

Meanwhile, many current visualization systems do not perform well in handling large amounts of data. We can combine visualization techniques with intelligent techniques, such as data mining and machine learning, to make effective detections.

Finally, we plan to develop a location service for local networks. The location service is meant to help administrators locate a specific IP address. As IP addresses are always assigned by an organization's structure, the hierarchical technology in [105] can briefly demonstrate the topology of departments but not the exact locations of each host. As in [123], a geography map is also provided as a plain in the three-plain integrated visualization system. However, it is hard to find an exact IP address in the integrated plains. Inspired by geolocation [104], we will establish an additional geography database in our design, based on which users can request the location information.

Table III. Summary of visualization designs.

Forms of visual results				
Text based	Parallel	Hierarchical	Three dimensional	Others
Geolocation [104], Log Parser [3], etc.	Picviz [116], Visual Firewall [119], Rumint [120].	Treemaps [122], IP-based hierarchical [105].	Inetvis [120], Integrated system [123], Flamingo [125], MineSet [97], A hybrid system [126].	BGP based [128], An e-mail monitoring system [132]

BGP, border gateway protocol.

DOI: 10.1002/sec

6. CONCLUSIONS

A visualization system provides a more perceptive method for security analysis. It is an effective way to use visual analytics in exploring logs, as this is easier for humans to find than unexpected patterns. Visual analytics are also advantageous for finding explanations or reasons for an observed phenomenon, especially for these open-ended problems. That is because visually explored data and patterns, together with the users' knowledge, can help users find explanations [2]. Automated algorithms and methods are more suitable to handle those problems with exhaustively enumerated reasons, which can be tested one by one [2].

In this paper, we studied visualization designs from five different categories, and in each category, we study several designs, as shown in Table III. For each design, we study its novel idea and detailed case with corresponding visual results. With the assistance of visualization systems, administrators can detect anomalies in a more perceptive and convenient way.

ACKNOWLEDGEMENT

This work is supported in part by the US National Science Foundation (NSF) under the grant numbers CNS-0737325, CNS-0716211, CCF-0829827, and CNS-1059265.

REFERENCES

1. Teoh ST, Ma KL, Wu SF. Visual exploration process for the analysis of internet routing data. In *Proceedings of the IEEE Conference on Visualization 2003*. IEEE Computer Society: Washington, DC, 2003; 523–530.
2. Teoh ST, Jankun-Kelly T, Ma K-L, Wu SF. Visual data analysis for detecting flaws and intruders in computer network systems. *IEEE Computer Graphics and Applications* 2004; **24**(5): 27–35.
3. Francia G, Trifas M, Brown D, Francia R, Scott C. Forensic Data Visualization System: Improving Security Through Automation. *Computer Security Conference*, Myrtle Beach, South Carolina, April 12-13, 2007.
4. Takahashi D, Xiao Y. Retrieving knowledge from auditing log files for computer and network forensics and accountability. *Security and Communication Networks* 2008; **1**(2): 147–160.
5. Xiao Y. Accountability for wireless LANs, ad hoc networks, and wireless mesh networks. *IEEE Communications Magazine*, Special Issue on Security in Mobile Ad Hoc and Sensor Networks 2008; **46**(4): 116–126.
6. Meng K, Xiao Y, Vrbsky SV. Building a wireless capturing tool for WiFi. *Security and Communication Networks* 2009; **2**(6): 654–668. DOI: 10.1002/sec.107
7. Xiao Y. Flow-net methodology for accountability in wireless networks. *IEEE Network* 2009; **23**(5): 30–37.
8. Zhuang Z, Li Y, Chen Z. Enhancing intrusion detection system with proximity information. *International Journal of Security and Networks* 2010; **5**(4): 207–219.
9. Abbes T, Bouhoula A, Rusinowitch M. Efficient decision tree for protocol analysis in intrusion detection. *International Journal of Security and Networks* 2010; **5**(4): 220–235.
10. Schrader KR, Mullins BE, Peterson GL, Mills RF. An FPGA-based system for tracking digital information transmitted via peer-to-peer protocols. *International Journal of Security and Networks* 2010; **5**(4): 236–247.
11. Chen Z, Chen C, Wang Q. On the scalability of delay-tolerant botnets. *International Journal of Security and Networks* 2010; **5**(4): 248–258.
12. Guo Y, Perreau S. Detect DDoS flooding attacks in mobile ad hoc networks. *International Journal of Security and Networks* 2010; **5**(4): 259–269.
13. Guo H, Mu Y, Zhang XY, Li ZJ. Enhanced McCullagh-Barreto identity-based key exchange protocols with master key forward security. *International Journal of Security and Networks* 2010; **5**(2/3): 173–187.
14. Richard AO, Ahmad A, Kiseon K. Security assessments of IEEE 802.15.4 standard based on X.805 framework. *International Journal of Security and Networks* 2010; **5**(2/3): 188–197.
15. Dong Y, Hsu S, Rajput S, Wu B. Experimental analysis of application-level intrusion detection algorithms. *International Journal of Security and Networks* 2010; **5**(2/3): 198–205.
16. Wang H, Jia X. Editorial. *International Journal of Security and Networks* 2010; **5**(2/3): 77–78.
17. Leng X, Lien Y, Mayes K, Markantonakis K. An RFID grouping proof protocol exploiting anti-collision algorithm for subgroup dividing. *International Journal of Security and Networks* 2010; **5**(2/3): 79–86.
18. Dalton II GC, Edge KS, Mills RF, Raines RA. Analysing security risks in computer and radio frequency identification (RFID) networks using attack and protection trees. *International Journal of Security and Networks* 2010; **5**(2/3): 87–95.
19. Mahinderjit-Singh M, Li X. Trust in RFID-enabled supply-chain Management. *International Journal of Security and Networks* 2010; **5**(2/3): 96–105.
20. Hutter M, Plos T, Feldhofer M. On the security of RFID devices against implementation attacks. *International Journal of Security and Networks* 2010; **5**(2/3): 106–118.

21. Imasaki Y, Zhang Y, Ji Y. Secure and efficient data transmission in RFID sensor networks. *International Journal of Security and Networks* 2010; **5**(2/3): 119–127.
22. Sun L. Security and privacy on low-cost radio frequency identification systems. *International Journal of Security and Networks* 2010; **5**(2/3): 128–134.
23. Zhang X, Gao Q, Saad MK. Looking at a class of RFID APs through GNY logic. *International Journal of Security and Networks* 2010; **5**(2/3): 135–146.
24. Azevedo SG, Ferreira JJ. Radio frequency identification: a case study of healthcare organisations. *International Journal of Security and Networks* 2010; **5**(2/3): 147–155.
25. Raad M. A ubiquitous mobile telemedicine system for the elderly using RFID. *International Journal of Security and Networks* 2010; **5**(2/3): 156–164.
26. Rodrigues MJ, James K. Perceived barriers to the widespread commercial use of radio frequency identification technology. *International Journal of Security and Networks* 2010; **5**(2/3): 165–172.
27. Yang M, Liu JCL, Tseng Y. Editorial. *International Journal of Security and Networks* 2010; **5**(1): 1–3.
28. Malliga S, Tamilarasi A. A backpressure technique for filtering spoofed traffic at upstream routers. *International Journal of Security and Networks* 2010; **5**(1): 3–14.
29. Huang S, Shieh S. Authentication and secret search mechanisms for RFID-aware wireless sensor networks. *International Journal of Security and Networks* 2010; **5**(1): 15–25.
30. Hsiao Y, Hwang R. An efficient secure data dissemination scheme for grid structure wireless sensor networks. *International Journal of Security and Networks* 2010; **5**(1): 26–34.
31. Xu L, Chen S, Huang X, Mu Y. Bloom filter based secure and anonymous DSR protocol in wireless ad hoc networks. *International Journal of Security and Networks* 2010; **5**(1): 35–44.
32. Tsai K, Hsu C, Wu T. Mutual anonymity protocol with integrity protection for mobile peer-to-peer networks. *International Journal of Security and Networks* 2010; **5**(1): 45–52.
33. Yang M. Lightweight authentication protocol for mobile RFID networks. *International Journal of Security and Networks* 2010; **5**(1): 53–62.
34. Wang J, Smith GL. A cross-layer authentication design for secure video transportation in wireless sensor network. *International Journal of Security and Networks* 2010; **5**(1): 63–76.
35. Bai L, Zou X. A proactive secret sharing scheme in matrix projection method. *International Journal of Security and Networks* 2009; **4**(4): 201–209.
36. Bettahar H, Alkubeily M, Bouabdallah A. TKS: a transition key management scheme for secure application level multicast. *International Journal of Security and Networks* 2009; **4**(4): 210–222.
37. Huang H, Kirchner H, Liu S, Wu W. Handling inheritance violation for secure interoperation of heterogeneous systems. *International Journal of Security and Networks* 2009; **4**(4): 223–233.
38. Rekhis S, Boudriga NA. Visibility: a novel concept for characterising provable network digital evidences. *International Journal of Security and Networks* 2009; **4**(4): 234–245.
39. Djenouri D, Bouamama M, Mahmoudi O. Black-hole-resistant ENADAIR-based routing protocol for mobile ad hoc networks. *International Journal of Security and Networks* 2009; **4**(4): 246–262.
40. Hu F, Dong D, Xiao Y. Attacks and countermeasures in multi-hop cognitive radio networks. *International Journal of Security and Networks* 2009; **4**(4): 263–271.
41. Chen Z, Chen C, Li Y. Deriving a closed-form expression for worm-scanning strategies. *International Journal of Security and Networks* 2009; **4**(3): 135–144.
42. Lee S, Sivalingam KM. An efficient one-time password authentication scheme using a smart card. *International Journal of Security and Networks* 2009; **4**(3): 145–152.
43. Watkins L, Beyah R, Corbett C. Using link RTT to passively detect unapproved wireless nodes. *International Journal of Security and Networks* 2009; **4**(3): 153–163.
44. Drakakis KE, Panagopoulos AD, Cottis PG. Overview of satellite communication networks security: introduction of EAP. *International Journal of Security and Networks* 2009; **4**(3): 164–170.
45. Chakrabarti S, Chandrasekhar S, Singhal M. An escrow-less identity-based group-key agreement protocol for dynamic peer groups. *International Journal of Security and Networks* 2009; **4**(3): 171–188.
46. Ehlert S, Rebahi Y, Magedanz T. Intrusion detection system for denial-of-service flooding attacks in SIP communication networks. *International Journal of Security and Networks* 2009; **4**(3): 189–200.
47. Berthier R, Cukier M. An evaluation of connection characteristics for separating network attacks. *International Journal of Security and Networks* 2009; **4**(1/2): 110–124.
48. Wu B, Wu J, Dong Y. An efficient group key management scheme for mobile ad hoc networks. *International Journal of Security and Networks* 2009; **4**(1/2): 125–134.

49. Mayrhofer R, Nyberg K, Kindberg T. Foreword. *International Journal of Security and Networks* 2009; **4**(1/2): 1–3.
50. Scannell A, Varshavsky A, LaMarca A, De Lara E. Proximity-based authentication of mobile devices. *International Journal of Security and Networks* 2009; **4**(1/2): 4–16.
51. Soriente C, Tsudik G, Uzun E. Secure pairing of interface constrained devices. *International Journal of Security and Networks* 2009; **4**(1/2): 17–26.
52. Buhan I, Boom B, Doumen J, Hartel, Veldhuis RNJ. Secure pairing with biometrics. *International Journal of Security and Networks* 2009; **4**(1/2): 27–42.
53. McCune JM, Perrig A, Reiter MK. Seeing-is-believing: using camera phones for human-verifiable authentication. *International Journal of Security and Networks* 2009; **4**(1/2): 43–56.
54. Goodrich MT, Sirivianos M, Solis J, Soriente C, Tsudik G, Uzun E. Using audio in secure device pairing. *International Journal of Security and Networks* 2009; **4**(1/2): 57–68.
55. Laur S, Pasini S. User-aided data authentication. *International Journal of Security and Networks* 2009; **4**(1/2): 69–86.
56. Suomalainen J, Valkonen J, Asokan N. Standards for security associations in personal networks: a comparative analysis. *International Journal of Security and Networks* 2009; **4**(1/2): 87–100.
57. Kuo C, Perrig A, Walker J. Designing user studies for security applications: a case study with wireless network configuration. *International Journal of Security and Networks* 2009; **409**(1/2): 101–109.
58. Ma L, Teymorian AY, Xing K, Du D. An one-way function based framework for pairwise key establishment in sensor networks. *International Journal of Security and Networks* 2008; **3**(4): 217–225.
59. Srinivasan A, Li F, Wu J, Li M. Clique-based group key assignment in wireless sensor networks. *International Journal of Security and Networks* 2008; **3**(4): 226–239.
60. Hsieh C, Chen J, Lin Y-B, *et al*. NTP-DownloadT: a conformance test tool for secured mobile download services. *International Journal of Security and Networks* 2008; **3**(4): 240–249.
61. Sadowitz M, Latifi S, Walker D. An iris and retina multimodal biometric system. *International Journal of Security and Networks* 2008; **3**(4): 250–257.
62. Kandikattu R, Jacob L. Secure hybrid routing with micro/macro-mobility handoff mechanisms for urban wireless mesh networks. *International Journal of Security and Networks* 2008; **3**(4): 258–274.
63. Xu H, Ayachit M, Reddyreddy A. Formal modelling and analysis of XML firewall for service-oriented systems. *International Journal of Security and Networks* 2008; **3**(3): 147–160.
64. Bouhoula A, Trabelsi Z, Barka E, Benelbahri M. Firewall filtering rules analysis for anomalies detection. *International Journal of Security and Networks* 2008; **3**(3): 161–172.
65. Li F, Srinivasan A, Wu J. PVFS: a probabilistic voting-based filtering scheme in wireless sensor networks. *International Journal of Security and Networks* 2008: **3**(3): 173–182.
66. Ma X, Cheng X. Verifying security protocols by knowledge analysis. *International Journal of Security and Networks* 2008; **3**(3): 183–192.
67. Uphoff B, Wong JS. An agent-based framework for intrusion detection alert verification and event correlation. *International Journal of Security and Networks* 2008; **3**(3): 193–200.
68. Tripathy S, Nandi S. Secure user-identification and key distribution scheme preserving anonymity. *International Journal of Security and Networks* 2008; **3**(3): 201–205.
69. Li F, Xin X, Hu Y. ID-based threshold proxy signcryption scheme from bilinear pairings. *International Journal of Security and Networks* 2008; **3**(3): 206–215.
70. Lin X, Ling X, Zhu H, Ho P, Shen X. A novel localised authentication scheme in IEEE 802.11 based wireless mesh networks. *International Journal of Security and Networks* 2008; **3**(2): 122–132.
71. Challal Y, Gharout S, Bouabdallah A, Bettahar H. Adaptive clustering for scalable key management in dynamic group communications. *International Journal of Security and Networks* 2008; **3**(2): 133–146.
72. Memon N, Goel R. Editorial. *International Journal of Security and Networks* 2008; **3**(2): 79.
73. Ray I, Poolsappasit N. Using mobile ad hoc networks to acquire digital evidence from remote autonomous agents. *International Journal of Security and Networks* 2008; **3**(2): 80–94.
74. Kilpatrick T, Gonzalez J, Chandia R, Papa M, Shenoi S. Forensic analysis of SCADA systems and networks. *International Journal of Security and Networks* 2008; **3**(2): 95–102.
75. Cronin E, Sherr M, Blaze M. On the (un)reliability of eavesdropping. *International Journal of Security and Networks* 2008; **3**(2): 103–113.
76. Okolica JS, Peterson GL, Mills RF. Using PLSI-U to detect insider threats by datamining e-mail. *International Journal of Security and Networks* 2008; **3**(2): 114–121.
77. Conti G. *Security Data Visualization. Graphical Techniques for Network Analysis*. No Starch Press: San Francisco, CA, 2007. ISBN-10 1-59327-143-3; ISBN-13 978-1-59327-143-5.

78. Clifton C. Encyclopedia Britannica: Definition of Data Mining. http://www.britannica.com/EBchecked/topic/1056150/data-mining, 2010.
79. Teoh ST, Ma K, Wu SF, *et al*. Visual-based anomaly detection for BGP origin as change (OASC) events. In *Proceedings of the Distributed Systems, Operations, and Management Workshop (DSOM'03)*. Springer: Heidelberg, Germany, Oct 2003; 155–168.
80. Fayyad U, Piatetsky-Shapiro G, Smyth P. From data mining to knowledge discovery in databases. *AI Magazine* 1996; **17**(3): 37–54.
81. Fayyad UM. Mining Databases: Towards Algorithms for Knowledge Discovery. *Data Engineering Bulletin* 1998; **21**(1): 39–48.
82. Fayyad UM, Piatetsky-Shapiro G, Smith P. From data mining to knowledge discovery: an overview. In *Advances in Knowledge Discovery and Data Mining*, Fayyad UM *et al*., (eds.) AAAI Press and MIT Press: Menlo Park, CA/Cambridge, MA, 1996; 1–34.
83. Schultz MG, Eskin E, Zadok E, Stolfo SJ. Data mining methods for detection of new malicious executables. In *Proceedings of the IEEE Symposium on Security and Privacy*, 2001.
84. Ghosh AK, Schwartzbard A. A study in using neural networks for anomaly and misuse detection. In *Proceedings of the Eighth USENIX Security Symposium*, 1999.
85. Jiang N, Hua K, Sheu S. Considering both intra-pattern and inter-pattern anomalies for intrusion detection, In *Proceedings of the 2002 IEEE International Conference on Data Mining (ICDM'02)*, 2002.
86. Mahoney MV, Chan PK. Learning models of network traffic for detecting novel attacks. In *Proceedings of the Third IEEE International Conference on Data Mining (ICDM'03)*, 2003.
87. Michael CC. Finding the vocabulary of program behavior data for anomaly detection. In *Proceedings DISCEX '03*, 2003.
88. Sun B, Yu F, Wu K, Xiao Y, Leung VCM. Enhancing security using mobility-based anomaly detection in cellular mobile networks. *IEEE Transactions on Vehicular Technology* 2006; **55**(4): 1385–1396.
89. Sun B, Wu K, Xiao Y, Wang R. Integration of mobility and intrusion detection for wireless ad hoc networks. *International Journal of Communication Systems* 2007; **20**(6): 695–721.
90. Sun B, Xiao Y, Wang R. Detection of fraudulent usage in wireless networks. *IEEE Transactions on Vehicular Technology* 2007: **56**(6): 3912–3923.
91. Herman I, Melançon G, Scott Marshall M. Graph visualization and navigation in information visualization: a survey. *IEEE Transactions on Visualization and Computer Graphics* 2000; **6**(1): 24–43.
92. Erbacher RF, Walker KL, Fincke DA. Intrusion and misuse detection in large-scale systems. *IEEE Computer Graphics and Applications* 2002; **22**(1): 38–48.
93. Yurcik W, Lakkaraju K, Barlow J, Rosendale J. A prototype tool for visual data mining of network traffic for intrusion detection. In *Proceedings of the ICDM Workshop on Data Mining for Computer Security (DMSEC'03)*, 2003.
94. Girardin L. An eye on network intruder-administrator shootouts. In *Proceedings of the Workshop on Intrusion Detection and Network Monitoring (ID'99)*. USENIX Assoc.: Berkeley, CA, 1999; 19–28.
95. Takada T, Koike H. Tudumi: information visualization system for monitoring and auditing computer logs. In *Proceedings of the Sixth International Conference on Information Visualization*, 2002.
96. Ankerst M, Ester M, Kriegel H-P. Towards an effective cooperation of the user and the computer for classification. In *Proceedings of the Sixth International Conference on Knowledge Discovery and Data Mining (KDD '00)*, 2000.
97. Brunk C, Kelly J, Kohavi R. MineSet: An Integrated System for Data Access, Visual Data Mining, and Analytical Data Mining. In *Proceedings of the Third Conference on Knowledge Discovery and Data Mining (KDD-97)*, Newport Beach, CA, August 1997; 135–138.
98. Takada T, Koike H. Mielog: a highly interactive visual log browser using information visualization and statistical analysis. In *Proceedings of LISA XVI Sixteenth Systems Administration Conference*. The USENIX Association: Berkeley, CA, 2002; 133–144.
99. Cheswick B, Burch H, Branigan S. Mapping and visualizing the internet. In *Proceedings of the 2000 USENIX Annual Techincal Conference*, 2000.
100. Munzner T. Exploring large graphs in 3D hyperbolic space. *IEEE Computer Graphics and Applications* 1998; **18**(4): 18–23.
101. Munzner T, Hoffman E, Claffy K, Fenner B. Visualizing the global topology of the MBone. In *Proceedings of the 1996 IEEE Symposium on Information Visualization*. IEEE Computer Society: Washington, DC, 1996; 85–92.
102. Labovitz C, Malan GR, Jahanian F. Internet routing instability. *IEEE/ACM Transactions on Networking* 1998; **6**(5): 515–528.
103. Goldstein J, Roth SF, Mattis J. A framework for knowledge-based, interactive data exploration. *Journal of Visual Languages and Computing* 1994; **5**: 339–363.

104. Geolocations. Available from: http://stoneblog.stonesoft.com/2009/07/smc-videos-geolocations/
105. Itoh T, Takakura H, Sawada A, Koyamada K. Hierarchical visualization of network intrusion detection data, *IEEE Computer Graphics and Applications* 2006; **26**(2): 40–47.
106. Teelink S, Erbacher RT. Improving the computer forensic analysis process through visualization. *Communications of the ACM* 2006; **49**: 71–75.
107. Jeong CY, Chang BH, Na JC. A survey on visualization for wireless security. In *Proceedings of the Fourth International Conference on Networked Computing and Advanced Information Management, NCM '08*, Vol. **1**, 2008; 129–132.
108. LaMarca A. Place lab: device positioning using radio beacons in the wild. In *Proceedings of the Third International Conference Pervasive Computing (Pervasive 05), LNCS 3468*. Springer: Berlin, 2005; 116–133.
109. Available from: http://www.microsoft.com/DownLoads/details.aspx?FamilyID=890cd06b-abf8-4c25-91b2-f8d975cf8c07&displaylang=en#Overview
110. Teoh ST, Jankun-Kelly T, Ma K-L, Wu SF. Visual data analysis for detecting flaws and intruders in computer network systems. *IEEE Computer Graphics and Applications*, 2004; **24**(5): 27–35.
111. Muelder C, Ma K, Bartoletti T. Interactive visualization for network and port scan detection. In *Proceedings of Visualization for Computer Security*, October 2006.
112. Teerlink S, Erbacher RF. Improving the computer forensic analysis process through visualization. *Communications of the ACM* 2006; **49**(2): 71–75.
113. Conti G, Abdullah K. Passive Visual fingerprinting of network attack tools. In *ACM Workshop on Visualization and Data Mining for Computer Security (VizSEC/DMSEC)*, ACM: New York, NY, 2004; 45–54.
114. McPherson J, Ma K-L, Krystosk P, Bartoletti T, Christensen M. PortVis: a tool for port-based detection of security events. In *VizSEC/DMSEC '04: Proceedings of the 2004 ACM Workshop Visualization and Data Mining for Computer Security*, 2004.
115. Bogen A, Dampier D, Carver J. Support for computer forensics examination planning with domain modeling: a report of one experiment trial. In *Proceedings of the 40th Hawaii International Conference on Systems Sciences*, January 2007; pp. 267b.
116. Tricaud S. Picviz: finding a needle in a haystack. In *Proceedings of the first UNIX Workshop on the Analysis of System Logs*, San Diego, CA, December 2008.
117. Inselberg A, Dimsdale B. Parallel coordinates: a tool for visualizing multi-dimensional geometry. In *VIS '90: Proceedings of the First Conference on Visualization '90*, Los Alamitos, CA. IEEE Computer Society Press: Washington, DC, 1990; 361–378.
118. Tricaud S. Picviz. http://archive.hack.lu/2008/picviz-hacklu2008.pdf
119. Lee CP, Trost J, Gibbs N, Raheem B, Copeland JA. Visual Firewall: real-time network security monitor. In *IEEE Workshop on Visualization for Computer Security*. IEEE Computer Society Press: Washington, DC, 2005; 129–136.
120. McRee R. Security visualization: what you don't see can hurt you. *Information Systems Security Association (ISSA) Journal*, June 2008.
121. Available from: http://en.wikipedia.org/wiki/Treemapping
122. Johnson B, Shneiderman B. Treemaps: a space-filling approach to the visualization of hierarchical information structures. In *Proceedings of the 2nd International IEEE Visualization Conference*, pp. 284–291, October 1991.
123. Mukosaka S, Koike H. Integrated visualization system for monitoring security in large-scale local area network. In *Proceedings of the Sixth International Asia-Pacific Symposium on Visualization, APVIS '07*. IEEE Computer Society: Washington, DC, 2007; 41–44.
124. Available from: http://flamingo.merit.edu/
125. Available from: http://secviz.org/content/flamingo-port-scan
126. Peng J, Feng C, Rozenblit JW. A hybrid intrusion detection and visualization system. *Proceedings of the 13th Annual IEEE International Symposium and Workshop on Engineering of Computer Based Systems*. IEEE Computer Society: Washington, DC, 2006; 505–506.
127. Hu F, Malkawi Y, Kumar S, Xiao Y. Vertical and horizontal synchronization services with outlier detection in underwater sensor networks. *Wireless Communications and Mobile* 2008; **8**(9): 1165–1181.
128. Teoh ST, Ma KL, Wu SF, Zhao X. Case study: interactive visualization for internet security. In *13th IEEE Visualization 2002 (VIS 2002)*, 2002.
129. Teoh ST, Zhang K, Tseng S, Ma KL, Wu SF, Lunt T. Detecting intruders in computer systems. In *Proceedings of the 1993 Conference on Auditing and Computer Technology*, 1993.
130. Lunt T. Detecting intruders in computer systems. In *Proceedings of the 1993 Conference on Auditing and Computer Technology*, 1993.
131. Rekhter Y, Li T. *A border gateway protocol 4 (bgp-4)*. RFC 1771, 1995.
132. Inappropriate Content Visualization. http://5thsentinel.wordpress.com/2009/04/01/inappropriate-content-visualization/

133. Fink GA, Duggirala V, Correa R, North C.Bridging the host-network divide: survey, taxonomy, and solution. In *Proceedings of LISA '06*, Washington, DC. USENIX Association: Berkeley, CA, 2006; 247–262.

134. Teoh ST, Ma KL, Wu SF, Zhao X. CluVis dual-domain visual exploration of cluster network metadata. In *Proceedings of the 45th Annual Southeast Regional Conference, ACMSE '07*, 2007.

135. Shneiderman B. *Designing the User Interface: Strategies for Effective Human-Computer Interaction: Second Edition*, Addison-Wesley Publ. Co.: Reading, MA, 1992.

136. Muelder C, Ma K-L, Bartoletti T. A visualization methodology for characterization of network scans. In *Proceedings of the IEEE Workshop Visualization for Computer Security (VizSEC)*, Oct. 2005.

137. Lee W, Stolfo SJ, Mok KW. A data mining framework for building intrusion detection models. In *Proceedings of the IEEE Symposium on Security and Privacy*, 1999.

138. Shneiderman B. *Designing the User Interface: Strategies for Effective Human–Computer Interaction: Second Edition*, Addison-Wesley Publ. Co: Reading, MA, 1992.

139. Ma K-L. Visualization for security. *Computer Graphics* 2004; **38**(4): 4–6.

140. Peter L, Varian HR. *How Much Information*, 2003. Available from: http://www.sims.berkeley.edu/how-much-info-2003