SCIENCE CHINA
Information Sciences

• REVIEW •

December 2017, Vol. 60 121101:1–121101:17
doi: 10.1007/s11432-016-0428-2

A survey of network anomaly visualization

Tianye ZHANG, Xumeng WANG, Zongzhuang LI, Fangzhou GUO, Yuxin MA & Wei CHEN*

State Key Laboratory of CAD & CG, Zhejiang University, Hangzhou 310058*, China*

Received July 7, 2016; accepted August 12, 2016; published online April 24, 2017

Abstract Network anomaly analysis is an emerging subtopic of network security. Network anomaly refers to the unusual behavior of network devices or suspicious network status. A number of intelligent visual tools are developed to enhance the ability of network security analysts in understanding the original data, ultimately solving network security problems. This paper surveys current progress and trends in network anomaly visualization. By providing an overview of network anomaly data, visualization tasks, and applications, we further elaborate on existing methods to depict various data features of network alerts, anomalous traffic, and attack patterns data. Directions for future studies are outlined at the end of this paper.

Keywords network anomaly, network anomaly visualization, visual analysis, network security, visualization

Citation Zhang T Y, Wang X M, Li Z Z, et al. A survey of network anomaly visualization. Sci China Inf Sci, 2017, 60(12): 121101, doi: 10.1007/s11432-016-0428-2

1 Introduction

In recent years, the rapid development of network communication techniques is increasingly exerting rigorous and scalable requirements on network security awareness [1]. Thus, analysis of network monitoring data plays a key role for both researchers and network service providers.

In the field of network data analysis, analyzing network anomaly data is one of the most significant challenges. Network anomaly refers to the unusual behavior of network devices or suspicious network status, which can either be malicious or benign and can be attributed to network attacks or failure of network devices, such as abnormal network flows, unauthorized access of resources, and suspicious host behavior [2]. Network attack brings an outbreak of vicious abnormal activities whereas network failure causes harmless abnormal behavior.

Analyzing network anomaly is of great significance and value. Understanding the difference between normal network events and anomalous activities is essential for monitoring network security, thereby suggesting the behavior of potential security risks and presenting network conditions. In this way, detecting and reasoning of network anomaly can ensure damage mitigation and future prevention of possible network attack or network failure. Understanding network anomaly can benefit network security analysis in an intuitive way.

* Corresponding author (email: chenwei@cad.zju.edu.cn)

Table 1 Classification of network anomaly data and relevant visualization tasks

Data type		Data properties	Visualization tasks		
			Detection & identification	Correlation & classification	Awareness & assessment
Network alerts		Temporal	[3]	[4–6]	[7, 8]
		Tabular	[9, 10]	[11, 12]	[13, 14]
		High-dimensional		[15]	[16]
		Topological	[17–20]		[21, 22]
Anomalous traffic	Host	Temporal	[23–26]		
		Tabular	[27–30]	[31]	
		Topological		[32, 33]	
	Port	Temporal	[34, 35]	[36]	
	Multi-level	Topological	[37–42]	[2, 43]	
		Tabular	[44, 45]		
Attack patterns		Tabular			[46]
		Topological			[47]
Others		Spatial	[48]	[49]	
		High-dimensional		[50]	

From the aspect of network security analysts, visualization tools are extensively adopted in a variety of areas such as real-time monitoring, offline log analysis, and attack detection. Visualization methods turn inflexible network anomaly data into graphical representations and provide interactive analysis modes, which completely enhance human perception. By leveraging statistical charts or specifically-designed visual summarization and representation, complex abnormal patterns from massive network datasets are visually presented and obtained by human analysts for further decision making.

This paper aims to fill the gap between network security analysis and information visualization and summarize the state-of-the-art network security visualization. First, we provide a summarization and categorization of data types and features in Section 2 with their corresponding analytical tasks and network security applications from the aspect of visualization in Section 3. In the following sections, we classify the existing work on network anomaly visualization based on the three categories of data types presented in Table 1, which includes network alerts (Section 4), anomalous traffics (Section 5) and attack patterns (Section 6). Finally, we identify the challenges for future research in Section 7.

2 Data

Different types of network anomaly data are used to obtain different purposes with regard to the network security community. Table 1 summarizes the most commonly used data types and their properties correlated with potential visualization tasks.

2.1 Network anomaly data

Network anomaly data refer to unusual network events and trends recorded in network traces or detected by automated network defense devices. The anomaly data that draw the most attention of the network security community can be roughly categorized into the following three categories [31, 43].

• **Network alerts**. Network defense devices, e.g., intrusion detection systems, firewalls, anti-virus software, and others, generate alert events by automating matching algorithms based on a fundamental network information, such as network status and traffic data [51]. Each generated alert contains at least three kinds of information, as follows: time stamp occurrence, alert type, and involved devices (always in the form of source and destination IP addresses).

• **Anomalous traffic**. The preceding conventional automated network security systems could fail to detect all potential anomalies and even occasionally mislead analysts to draw erroneous conclusions [51].

Table 2 Visualization forms for different data features

Data features & visualization forms	Tabular(%)	Temporal(%)	Topological(%)	Spatial(%)	High-dimensional(%)
Statistical chart (bar/line chart, etc.)	22.58	46.43	–	–	–
Node-link diagram	16.13	–	68.75	–	50
Parallel coordinates	9.68	–	–	–	–
Matrix/grid	29.03	21.43	–	–	25
Scatter plot	12.90	14.28	–	–	–
Stack graph	–	10.71	–	–	–
Radial layout	3.23	7.14	18.75	–	–
Bundle diagram	3.23	–	6.25	–	–
Tree map	–	–	–	–	25
Hierarchical tree	3.23	–	6.25	–	–
Physical map	–	–	–	100	–

Directly analyzing anomalous network traffic with raw packets or network flow records effectively addresses this problem by incorporating human perception as well as expertise. Unlike network alert events, this type of anomaly data is hidden in normal traffic events and is identified with the help of both visualization methods and human knowledge. Based on the distinguished analysis level of network traffic, anomalous traffic can be further categorized into three classes, as follows: traffic of hosts, traffic of ports, and multi-level traffic.

• **Attack patterns**. Network attack is a more specifically analyzed object of network security compared with alert events and anomalous traffic, involving a sequence of devices that serve as a way to reach the ultimate goal of the attacker [47]. The security community is therefore quite generous for focusing their attention on the attack paths and the respective patterns of complex attacks.

In addition to the preceding three data types, other network anomaly data can also be derived from network information gathering facilities, such as malware analysis tools and compound results of mathematical computing models; for example, previous studies [52–54] focused on single or multiple features of malware samples to understand the in-depth exploration of malware behavior. Matuszak et al. [49] calculated network trust by using a mixture of different proportions of availability, detection, and false alarm values, as well as predictability. Kotenko et al. [50] computed network vulnerability based on specific scoring systems.

2.2 Data properties

As previously stated, most network anomaly data are originally derived from system logs generated by security maintaining systems or raw network trace records. This particular kind of data source allows the coexistence of a certain number of data features (as shown in Table 1), in which each demands different visualization and analysis methods. Table 2 is a summarization of common visualization forms for different data features.

• **Tabular**. Logs and records always appear in the form of tabular format that can be considered as a series of network events records.

• **Temporal**. Each record in the entire series contains a time stamp to determine when a message is generated.

• **Topological**. The records in the series also contain relation and information transfer between devices that provides the availability for the entire network topology.

• **Spatial**. In a few cases, the records in the series may contain spatial property such as locations to indicate the physical position of each network event, which is complementary to the logical position in network topology.

• **High-dimensional**. Apart from the previously stated time and location properties, the records in the series are also likely to include other information, such as relevant ports, transport protocols, and so on. Therefore, representing each record as a high-dimensional vector is natural.

Table 3 Presented data sources in available materials

Data source		Tools	Dataset
Security events	Firewall	Lee et al. [7]	–
		Foresti et al. [6]	IC-ARDA research projects.
	IDS	Shiravi et al. [11]	Private Snort dataset.
		Abdullah et al. [13]	Georgia tech campus alarm logs.
		Girardin et al. [15]	IES project.
		Koike et al. [4]	Signature database of NIDS.
		Onut et al. [27]	DARPA datasets.
Network traces	Netflow records	Ren et al. [23]	DNS query logs from a US university.
		Yin et al. [25]	Cisco.
		Taylor et al. [37]	SiLK.
		Mcpherson et al. [35]	Department of energy traffic.
	Packets	Ball et al. [28]	TCP-dump & Ethereal.
		Abdullah et al. [36]	TCP-dump and Pcap.
Other	Malware	Inoue et al. [44]	Malware knowledge pool.

2.3 Data sources

Different data sources include different kinds of network anomaly data with various data features. Table 3 is a summarization of presented data sources in the available materials. The sources of security events include network alert and attack pattern data, whereas sources of network traces include anomalous traffic data.

These datasets vary in format, but all contain crucial information that plays a significant role in visualization and analysis, such as time stamp, type of protocols (e.g., TCP, UDP, and others), type of service (e.g., FTP, DNS, Web, and others), source and destination of IP address and port, and detailed information of a certain event. These data contribute to the foundation of the analysis system and the analysis process.

In most cases, visualization systems are based on a single data source to avoid the problem of the time-consuming heterogeneous data processing, whereas incorporating several data sources into a single system allows analysts to obtain a richer insight by integrating all kinds of data and data features.

3 Visualization tasks, applications and pipeline

Visualization methods are extensively applied in the analysis process of network anomaly data to support the needs of network security.

3.1 Tasks of visualization

The tasks of visualization analysis, which includes detection and identification, correlation and classification, as well as awareness and assessment, are designed accordingly in order to fulfill the various requirements of network anomaly research.

• **Detection and identification**. Network anomaly analysis cannot be conducted without targets. The first step of anomaly analysis is to identify all suspicious anomalous events, thereby earning visualization tasks of anomaly detection and identification an essential position. Instead of only displaying anomaly generated by automated detection methods, visualization tasks require graphical representations of fundamental network information, such as network traffic, to help analysts distinguish anomalous network behaviors from a substantial volume of normal status [51].

• **Correlation and classification**. Network anomalies vary in forms and involve numerous devices (or network entries of a certain device). Classifying these two aspects and discovering correlating patterns between different classification groups through visualization methods are actually necessary. In this way, security operators are capable of determining more harmful anomaly groups that cause greater damage

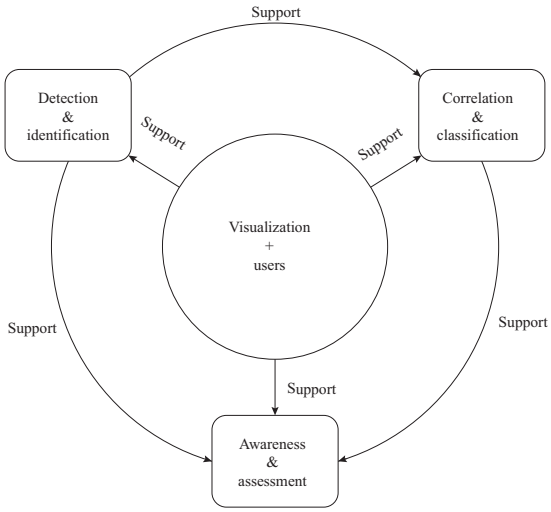

Figure 1 Partnerships between user-involved visualization and visualization tasks.

and locate more vulnerable device groups, which are easily destroyed by a specific kind of attack. Such visualization tasks require displaying information of different groups together in one or multiple views to assist with comparison and correlation.

• **Awareness and assessment**. Another kind of visualization aimed at raising the overall awareness of network anomalous state, including the temporal distribution and the varying trend. Therefore, assessment of a specific anomaly can be created based on its severity, frequency, and effects. Such visualization tasks generally require visual displays of multi levels and multi-dimensions to provide an overview-to-detail analysis mode.

These tasks are all supported by user-involved visualization, as shown in Figure 1. The tasks of raising awareness and creating an assessment is a relatively more comprehensive goal of network anomaly visualization and is occasionally performed with the assistance of detection results as well as correlation and classification analysis. Moreover, the task of detecting and identifying anomaly can also provide the foundation for correlation and classification. A complex visualization system may be designed to fulfill two or more tasks among these three in order to support a multi-stage analysis process.

3.2 Applications of visualization

Three visualization tasks provide different levels of security maintaining purposes, thereby serving for the same goal that is applied to specific security use cases. Visualization of all kinds of network anomaly data can be applied to analyze various network security problems with an attempt to alleviate or even solve the problem.

For example, Lee et al. [7] present visual firewall that accurately depicts patterns of distributed denial of service attack by visualizing network alerts of intrusion detection system logs and other network information altogether.

SVision [27] uses anomalous network traffic recorded in packet trace to distinguish the victims and the attacker of a Nachi worm attack. The goal of SVision is to integrate visualization with human knowledge to develop beneficial means for security operators to ease the tedious process of mitigating the worm.

Yelizarov and Gamayunov [46] design a visualization tool to display complex attacks. Therefore, complicated security problems such as distributed network scanning can be analyzed in a more interactive and understandable way.

In addition to the three preceding security problems (network scanning, DOS attack and worm attack), other application scenarios are also seen in network anomaly visualization. Table 4 provides a rough summarization of common security applications of each type of anomaly data.

Table 4 Different applications of all kinds of network anomaly data, including network alerts, anomalous traffic, attach patterns, and other data

Applications	Network alerts visualization	Anomalous traffic visualization	Attack patterns visualization	Other visualization
Port and network scanning	[3, 6, 7, 9, 11, 15, 16]	[9, 24, 28, 30, 33–37, 45]	[12, 46, 47]	[55]
DOS attack	[6, 7, 18, 20]	[25, 27, 34, 37, 43]	[2, 46]	
Viruses/Worms/Trojans attack	[4, 6–8, 13, 20, 36]	[25, 27, 29]	[2]	
IP spoofing/swapping	[13, 15]	[37]	[14]	
Backdoors and rootkits	[9, 18]	[36, 37]		
Propagate spam	[5, 8, 19, 21]			
Misuse of computer networks		[29, 30, 44]	[22]	
Zombie networks		[23]		
Malware infections		[31, 33, 43]		[50]

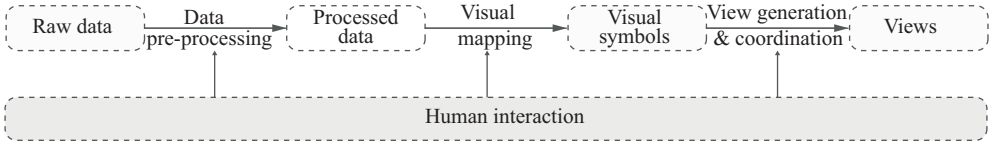

Figure 2 Conceptual pipeline of network anomaly visualization.

3.3 Pipeline of visualization

Visualization of network anomaly data usually follows a general visualization pipeline [56], as depicted in Figure 2.

Raw data derived from systems logs or network records may contain a variety of errors and invalid items; therefore, employing a series of data pre-processing operations, including data cleaning, data aggregation, data transformation, and others is necessary. These operations not only remove the redundancy of raw data but also process data into a structured format for the next stages.

Once the datasets are ready, the visual mapping is applied to convert data objects into visual symbols. Visual symbols with targeted designs are required to adapt to the need of displaying different data features for network alert, traffic, and attack data. Effective visual designs ensure the participation of human perception and enhance the ability to obtain insights.

Finally, the visual symbols are integrated into single or multiple views to fulfill analysis demands. Meanwhile, interaction is available during each stage. Typical interactions include choosing, dragging, zooming, adjusting color mapping and data mapping ways, and level-of-detail control. Intelligible and intuitive interactions allow analysts to participate in the entire visualization pipeline and make the biggest use of their expertise by adjusting the parameters of data processing procedures, selecting different visual mapping methods, or manipulating visual representatives. The entire pipeline and the feedback loop increase the possibility of uncovering hidden patterns in the data and generating beneficial knowledge of the target network system.

4 Network alerts visualization

Visualization of network alert data are mostly aimed at fulfilling tasks of correlation and classification, as well as awareness and assessment. Multiple data features (as illustrated in Section 2) are used to support this procedure. This section describes the visualization techniques specially designed for temporal data, tabular data, high-dimensional data and topological data in network alerts.

4.1 Temporal data

With regard to time periodicity, temporal data can be classified into linear time and cyclic time [57]. Temporal feature visualization of network alerts allows analysts to discover the anomalous trend, anomaly

occurrence frequency as well as periodicity, and others.

4.1.1 *Linear time*

Linear time uses a linear time domain to interpret time elements from a starting to an ending point [57]. The linear feature of linear time perfectly suits the visual representation of elements along a linear axis [58–69].

For example, SnortView represents time along the horizontal axis and aligns the source IP along the vertical axis. The entire coordinate system is therefore divided into numerous small cells in which multiple glyphs of varying shapes and colors depict different services and protocols of NIDS alarms. Different from SnortView, Foresti et al. projected time to a linear axis along the radius direction and proposed the W^3 visualization concept that stands for where, when, and what. A node-link diagram in the inner circle indicates the entire network topology, which is linked with dark dots on the outer rings to represent different types of network alerts. In this way, correlations between various alerts can be identified by filtering alerts and investigating different time stages. However, only alerts occurring in the latest history period are shown in the system to avoid visual clutter attributed to complex situations. As a result, this visualization method may work well with a short time period but is not appropriate for long time periods.

Another effective visualization form for linear time is stack graph, which displays multiple values of different aspects of a specific time stamp. EvoRiver [70] and CiteRivers [71] used a stack graph-based theme river to show topic and citation coopetition dynamics during a certain linear time period. Unfortunately, stack graph is rarely seen in network anomaly visualization in spite of its excellence.

4.1.2 *Cyclic time*

Cyclic time refers to time variables that vary periodically (for example, 24 hours of each day [72], four seasons of each year). Radial visualization designs are obviously more appropriate for a cyclic time.

SpiralView [5] applies a spiral time axis to highlight the periodical alerts. In this layout, a circle represents a day, and each circle is divided into 24 parts depicting 24 hours. Therefore, analysts can easily notice the periodical alerts that last longer than the others. However, although radial layouts clearly display periodic patterns, they cannot avoid the problem of space wasting [73].

Visualization methods designed for cyclic time are more suitable for periodically regular data with the aim of discovering abnormal patterns. In contrast, analysts benefit from linear time representations that appear more intuitive in indicating attribute values. Deviant values and sudden changes of attributes are more obvious in a linear time display.

4.2 Tabular data

As previously stated, alert data are always recorded as a series of events in network logs, thereby illustrating what is happening to the target network. Based on the aggregation level of network alerts, visualization of tabular alert data can be categorized into a single event based visualization and group based visualization.

4.2.1 *Single event based visualization*

Single event based visualization considers recorded events of network alerts as separate individuals and uses consistent visual elements to present these events one after another. Popular visualization forms are the list, scatterplot, pixel map, and so on.

For example, Lamagna proposed a pixel-map-based heat map design to visualize network log alarms [9], as shown in Figure 3. Each block in the heat map represents a record in a log file, arranged in a day (Y axis) to hour (X axis) way with varying colors indicating different sources of logs files. Another example of pixel map design is the system developed by Nicklaus et al. [10], which include a bivariate geomap view, a scatterplot view, a parallel coordinate plots view, and a histogram view. Alert frequency and types are clearly described in this design.

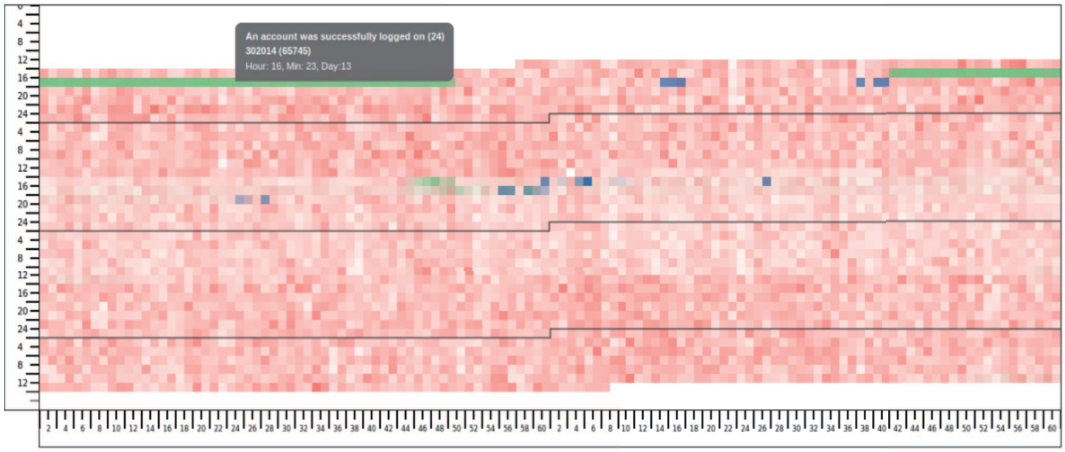

Figure 3 (Color online) An integrated visualization on network events [9] @ copyright 2011 IEEE.

Single event based visualization in most cases involves a massive number of alerts. The advantage of a single event based visualization is that it offers an overview of all recorded network alerts and shows the distribution pattern of alerts in the entire event series.

4.2.2 *Group based visualization*

Group based visualization techniques are designed to display network alerts in aggregated groups. Alert types and network attributes are commonly used as classification rules.

IDS RainStorm [13] is an example of supervising the alerts of departmental networks by aggregating data based on both IP address and time to reduce overlaps.

Shiravi et al. [11] implemented Avisa, a security visualization system that is aimed at enhancing the capability of analysts to correlate different types of alerts. IDS alerts are categorized into groups at the top left corner and the curves connect relevant hosts to corresponding alert types. By correlating prioritized hosts to intrusion alerts, Avisa makes it possible for analysts to compare hosts confronted with different network intrusions and distinguish which intrusion types deserve the main concern.

Comparing groups and determining correlations between groups can be achieved by group-based visualization. Analysts can identify the alert type or the vulnerable network and prioritize which should be handled first.

4.3 High-dimensional data

The coordinate system is suitable for two or three-dimensional data by projecting each attribute to each axis [74–76]; however, it can hardly handle high-dimensional data with four or more dimensions. Visualization of high-dimensional data aims at displaying high-dimensional data in low-dimensional spaces (always in two-dimensional spaces). Common techniques are icon displaying and spatial mapping [77].

4.3.1 *Icon displaying*

Specifically designed icons, which use various visual elements of the icon to represent different attributes of the data object, are popular in displaying high-dimensional data.

Radar chart (or star plot) is an typical example of these icons [68,78,79]. Data attributes are valued by axes radially arranged along the circle sharing the same starting point and lines connecting neighboring attributes to show the overall state of data. Ref. [78] summarizes anomalous communication features of network users.

Radar chart provides a compact visualization form to display high-dimensional data; however, this chart is inappropriate when the dataset becomes larger in volume rather than dimension.

4.3.2 *Spatial mapping*

Fortunately, spatial mapping methods can cover the drawback of icons. Spatial mapping methods use two-dimensional spatial location to display attributes as well as correlations between substantial amounts of high-dimensional data.

Parallel coordinates [80–85] and scatterplot [86–88] matrix are two of the extensively used spatial mapping techniques. Each axis in parallel coordinates represents a single attribute. Lines passing through all axes correspond to data objects. Scatterplot matrix visualized N-dimensional data with N^2 scatterplots in an $N \times N$ matrix, with each one indicating the correlation between a certain pair of attributes.

However, all the preceding visualization techniques may lose their efficacy once the data dimension becomes extremely high. In such cases, dimension reduction methods like principal component analysis (PCA) [89,90] and multidimensional scaling (MDS) [91–93] are required to project high-dimensional data into a low-dimensional space by either linear or nonlinear transformation.

NIVA [16] projects the four segments of IP address into three-dimensional Cartesian coordinate system by IP-Space algorithm based on MDS. In addition, a dimension reduction method for visualization that is called self-organizing map (SOM) algorithm is applied to the system proposed by Girardin [15]. SOM, which was invented by Kohonen [94], is a technique for reducing dimensions and displaying similarities and dissimilarities. The main difference between MDS and SOM is that classical MDS achieves the goals via eigenvalue decomposition whereas SOM is a kind of unsupervised learning algorithm.

Unlike icon-based techniques whose efficacy lies in presenting different attributes of single objects, spatial mapping techniques are capable of not only revealing the distribution characteristics of multiple objects in a single view but also effectively explaining the relationship between two attributes. As a result, spatial mapping techniques are always expected to provide a big picture of a dataset.

4.4 Topological data

Visualization of topological alert data displays alerts based on network structure. Employing a node-link diagram is a common technique to achieve this goal, in which each node represents an anomaly object and links connecting pairs of nodes to indicate some kind of relationship.

Maltego [17] used an optimized node-link layout to display clustering of online and network information to reveal potential events that threaten the operation environments of a certain organization. Ocelot [21] is a novel hybrid hierarchical node-link visualization that employs a packing technique called Petri dish, as shown in Figure 4. A convenient network visualization system can always meet the requirements of filtering, searching, selecting, and other interactions to facilitate the exploration of the network. Petri dish enables users to not only explore the topological structure but also freely group entities by selecting nodes. Hence, Ocelot can be applied to detect the anomaly and flexibly create related responses.

5 Anomalous network traffic visualization

Visualization of anomalous network traffic are mostly designed for tasks of detecting and identifying anomalous behaviors. This section introduces visualization techniques designed for temporal, tabular, and topological features of anomalous traffic data.

5.1 Temporal data

Visualization designs tend to focus on the varying trends of linear time rather than periodicity with regard to anomalous traffic data. The following examples introduce both conventional and novel ways of displaying linear time of traffic data.

Hosts serve as access to the network. When an anomaly occurs, host traffic immediately becomes strange, thereby making host traffic visualization significant. Ren et al. [23] presented a visualization system depicting dynamic DNS traffic on the host level by employing animation methods. Zhou et al. [26] proposed and implemented ENTVis to extend the entropy-based traffic analysis from the temporal space

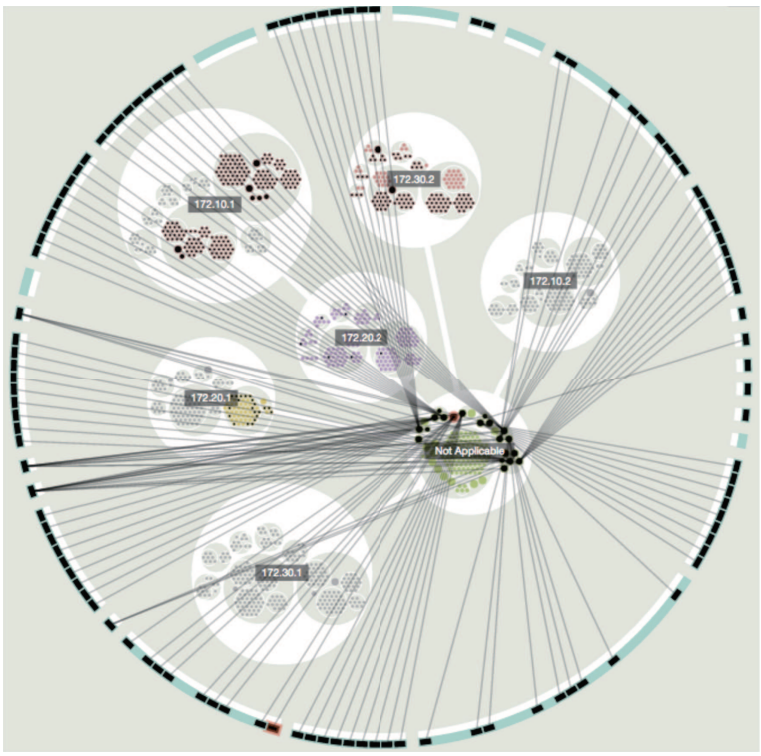

Figure 4 (Color online) The Petri dish of Ocelot [21] @ copyright 2015 IEEE.

to the visual clustering space, thereby quickly identify linear time spans when similar network traffic features occurred. TNV [24] lists time along the horizontal direction of the matrix and hosts along the vertical direction. One of the main ideas of TNV is to display host activities and their connections over time. Interactions, such as selecting areas of interest and filtering links, facilitate the exploration of network traffic to support anomalous behavior detection process. Provided with a certain time period and suspicious hosts, analysts are able to deduce problem hosts (both internal and external) involved in the attack based on the discovered prolonged packets sending strange-looking activities and inactive host behavior.

Ports are used as notations to identify different network processes of a single host. Port traffic serving as data transmission entries offers another anomaly detection mode similar to host traffic. Abdullah et al. [36] employed stacked histograms with horizontal time axes to visualize the number of packets delivered from different ports by various protocols. Similarly, IDGraphs [34] and PortVis [35] display port traffic data in a 2D space that aligns linear time information in the vertical or horizontal direction and another attribute in the opposite direction. Conventional statistical graphs such as these offer an entry to the analysis process by providing an overview of the entire traffic situation, thereby allowing exploration of temporal similarities and patterns. In this way, analysts can immediately identify and investigate suspicious ports in the port detail view, which interpret all available information of the selected port.

5.2 Tabular data

Different from network alert data, emphasizing the network traffic that flows from one device to another is necessary for visualization of tabular traffic data.

5.2.1 *Single event based visualization*

For single event based visualization of traffic data, directed lines and adjacency matrix are popular ways to indicate both the source and the destination of a traffic record.

VISUAL [28] and NICTER [44] are typical examples that display information transfer between hosts

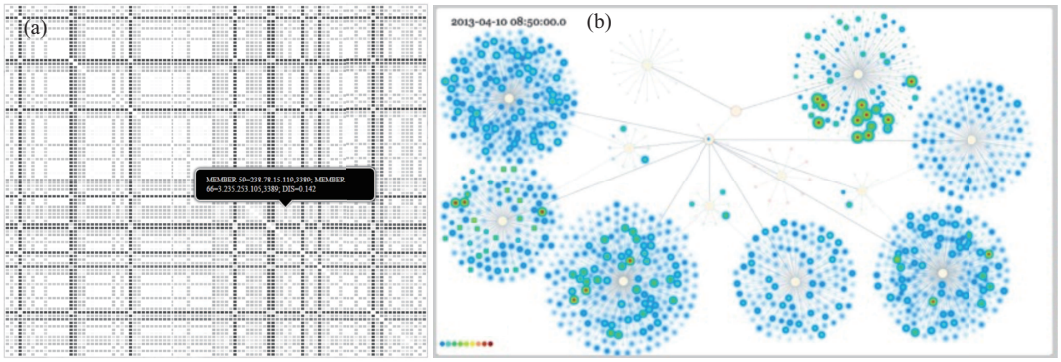

Figure 5 (Color online) Visualization design for tabular and topological traffic data. (a) Dissimilarity matrix designed by Hao et al. [31] @ copyright 2015 IEEE; (b) a node-link-based heat map design [39] @ copyright 2014 IEEE.

by directed lines. Distinguished spatial positions, such as inner and outer space as well as left and right space, are applied to differentiate the source devices from the destination devices.

Intuitively detecting an anomaly of hosts at a certain time point it not easy because of such visualization forms; however, these forms can hardly depict the dynamic change of traffic of a time period. Animation design is an alternative choice to improve these forms; however, brings memorial burden.

5.2.2 *Group based visualization*

Group based visualization of traffic data aims at comparing various groups of traffic to identify the outliers. Hao et al. [31] defined the dissimilarity of flow ensembles with regard to several aspects such as duration, density, and distribution. Figure 5(a) shows the dissimilarity matrix in which a darker color refers to the ensembles with higher dissimilarity. This visualization design allows analysts to focus on a small and manageable subset of network flow for follow-up analysis.

5.3 Topological data

The network topology is significant in visualization of traffic data. Analysts can easily identify groups of hosts that continuously communicate with each other, as well as the specific host that plays a central role by analyzing the topological structure.

DAEDALUS-VIZ [43] offers a highly flexible and tangible visualization system for darknet traffic. DAEDALUS-VIZ also employs animation to exhibit traffic events in real-time and 3D visualization to illustrate the complicated topological structure. As shown in Figure 6 in Ref. [43], the sphere in the center represents a complete IPv4 address space and the surrounding rings represent the organization under the surveillance of DAEDALUS. These lines and points between the sphere and the rings show the transmission of darknet packets. Furthermore, a variety of filters can be applied in this system, thereby allowing users to find significant information in a short time.

3D visualization is proven to be effective in facilitating the completion of complex tasks [95]. However, the basic interactions of 3D view always include 3D navigation, which is recognized as a kind of time-consuming and inaccurate interaction [96]. 2D visualization systems with appropriate designs with the purpose of visualizing anomalous network traffic are preferred by the visualization community.

Based on conventional node-link diagrams, Liao et al. [32] improved the node design by presenting informative nodes that include various anomaly events. Cortese et al. [41] enhanced the simple node-link BGPlay system by employing a topographic map that clearly shows ASes traversing tiers. Zhao et al. [39] developed a node-link-based heat map design for assessing global network security situation as shown in Figure 5(b), leading to a gateway of displaying the real-time active extent of host activities based on network topology. Locate active hosts and discovering subnet structures, together with long time observation that provides the possibility of detecting abnormal host activities, is quite convenient for analysts because of combined heat map and node-link diagram.

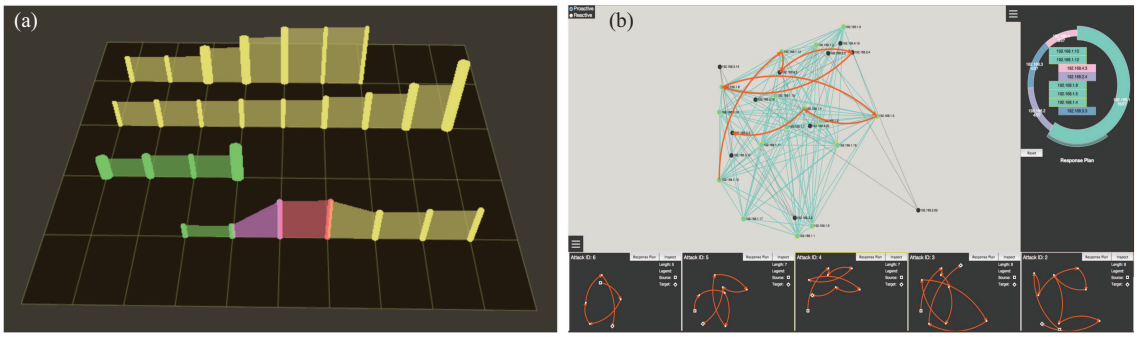

Figure 6 (Color online) Visualization design for network attack patterns. (a) Displaying multistep attacks in Ref. [46] @ copyright 2009 IEEE; (b) the PERCIVAL system relying on attack graph [47] @ copyright 2015 IEEE.

Glanfield et al. [38] created a novel network hierarchy presentation to better analyze network flow by aggregating IP addresses into groups as elements of the hierarchy. Concentric circles are used to indicate the branch depth of the hierarchy and identify the difference in traffic volume between different levels of this branch. Analysts are supposed to first locate the abnormal IP hierarchy through the ring-like overview and then drill down to the detailed IP and volume information by interactively selecting the target hierarchy. Mansmann et al. [42] compared three space-filling layouts in displaying network address hierarchy. Therefore, he concluded that the one-dimensional HistoMap is preferable at the AS level because of its scalability and stability; by contrast, the strip treemaps better fit the IP level analysis because they maintain the input order of nodes.

However, visualizing intricate traffic in the network topology is challenging because of the tremendous amount of traffic generated by network devices at any time. Techniques such as edge bundling [40, 97], graph compression [32], and clustering [98] are used to simplify the graph result to solve the complexity of the topology.

6 Network attack patterns visualization

Visualization of network attack patterns handles the task of obtaining awareness of attack components and assessing the evolution pattern of a certain attack. Improving existing network protection measures or even inventing more consolidated ones against the analyzed attack to build a safer network system is possible for security operators because of this kind of visualization. Visualizing the attack path is the core of this kind of visualization.

Yelizarov and Gamayunov [46] used cylinder glyph connected by transparent quadrangles to indicate attacks at different stages during a complex attack. Varying sizes and colors of each cylinder depict the severity level and the attack type, respectively, as shown in Figure 6(a).

Angelini et al. presented PERCIVAL [47], a visual analytics tool that assists analysts in obtaining a better understanding of both static and dynamic risk levels of the network by applying proactive and reactive modes of the prototype based on the attack graph, as shown in Figure 6(b). Each node of the attack graph represents an attack step toward a certain device, in which the links between nodes indicate the transitions between attack steps. In the proactive mode, when the system is not currently under attack, comparison of computed attack paths analysis and the actual evolution of attack is allowed when the system is not currently under attack, thereby providing insights on the effectiveness of the adopted mitigation actions. However, the reactive mode is triggered whenever an attack is detected. The barriers of mitigation actions, which infer whether the action was successfully completed or not, are placed right beside the related nodes with different colors.

The attack graph based on the node-link diagram applied in the system displays important links (attack steps) well based on the entire network topology but fails to solve the visual occlusion problem when the graph becomes more complicated. Another advantage of PERCIVAL is the the use of multiple coordinated views. Attacks and the global environment of the network can be dynamic and extremely complicated,

which requires the integration of multiple visualization methods to support various levels of investigation. In such cases, association analysis of multiple views is necessary, as presented in Ref. [99–108].

7 Conclusion and future work

Visualization of network anomaly data provides analysts with interactive visual tools to integrate human perception and knowledge into the analysis process of network security. Visualization also addresses the problems arising from large-scale and heterogeneous data. This paper provides an overview of relevant visualization techniques and visual analysis systems with regard to network anomaly data. Fruitful researches have already been conducted in this field. Nevertheless, studies that have hardly been done before but are of great value still exist:

• **Real-time data analysis**. Most existing network anomaly visualization uses data from offline datasets. However, real-time analysis is actually worth more attention because of the support of timely response and decision making. Considering the large volume of real-time network anomaly data to be addressed, meeting the requirements of real-time display and response while remaining functional is quite challenging for visualization.

• **Big screen technique**. The amount as well as the content of network anomaly data are increasingly becoming rich in recent years because of the gradually development of malicious network actions. As a result, the need for simultaneously displaying more network information from different aspects arises. Under the leading trend of big screen employment, providing big screen technique an opportunity is worth trying. However, appropriately displaying different network information on each screen is still a challenge. Analysts could easily become confused when considering too much information.

• **Privacy**. Network security data, especially network anomaly data, are expected to maintain privacy and security whereas visualization designs require detailed data information to fulfill analysis tasks. Problems in balancing privacy needs of anomaly data and visualization demands should be solved.

• **Objective verification**. Verifications of existing network anomaly visualization systems are mostly subjective and case-dependent. Therefore, building a consistent and unified verification system that applies to any possible situation is difficult. This difficulty is holding back the urgent necessity of creating such an authoritative objective verification system.

Acknowledgements This work was supported by National Basic Research Program of China (973 Program) (Grant No. 2015CB352503), Major Program of National Natural Science Foundation of China (Grant No. 61232012), National Natural Science Foundation of China (Grant Nos. 61422211, u1536118, u1536119), Zhejiang Provincial Natural Science Foundation of China (Grant No. LR13F020001), and Fundamental Research Funds for the Central Universities.

Conflict of interest The authors declare that they have no conflict of interest.

References

1 Shiravi H, Shiravi A, Ghorbani A A. A survey of visualization systems for network security. IEEE Trans Visual Comput Graph, 2012, 18: 1313–1329
2 Pearlman J, Rheingans P. Visualizing network security events using compound glyphs from a service-oriented perspective. In: Proceedings of the Workshop on Visualization for Computer Security, Sacramento, 2008. 131–146
3 Janies J. Existence plots: a low-resolution time series for port behavior analysis. In: Proceedings of the 5th International Workshop on Visualization for Computer Security, Cambridge, 2008. 161–168
4 Koike H, Ohno K. SnortView: visualization system of snort logs. In: Proceedings of the 2004 ACM Workshop on Visualization and Data Mining for Computer Security, Washington, 2004. 143–147
5 Bertini E, Hertzog P, Lalanne D. Spiralview: towards security policies assessment through visual correlation of network resources with evolution of alarms. In: Proceedings of IEEE Symposium on Visual Analytics Science and Technology, Washington, 2007. 139–146
6 Foresti S, Agutter J, Livnat Y, et al. Visual correlation of network alerts. IEEE Comput Graph, 2006, 26: 48–59
7 Lee C P, Tros J, Gibbs N, et al. Visual firewall: real-time network security monitor. In: Proceedings of IEEE Workshop on Visualization for Computer Security, Minneapolis, 2005. 129–136

8 Koike H, Ohno K, Koizumi K. Visualizing cyber attacks using ip matrix. In: Proceedings of IEEE Workshop on Visualization for Computer Security, Minneapolis, 2005. 91–98
9 Lamagna W M. An integrated visualization on network events vast 2011 mini challenge #2 award: outstanding integrated overview display. In: Proceedings of IEEE Conference on Visual Analytics Science and Technology, Providence, 2011. 319–321
10 Giacobe N A, Xu S. Geovisual analytics for cyber security: adopting the geoviz toolkit. In: Proceedings of IEEE Conference on Visual Analytics Science and Technology, Providence, 2011. 315–316
11 Shiravi H, Shiravi A, Ghorbani A A. IDS alert visualization and monitoring through heuristic host selection. In: Proceedings of International Conference on Information and Communications Security, Barcelona, 2010. 445–458
12 Erbacher R F. Intrusion behavior detection through visualization. In: Proceedings of IEEE International Conference on Systems, Man and Cybernetics, Washington, 2003. 2507–2513
13 Abdullah K, Lee C, Conti G, et al. IDS rainstorm: visualizing IDS alarms. In: Proceedings of the IEEE Workshops on Visualization for Computer Security, Minneapolis, 2005. 1
14 Erbacher R F, Walker K L, Frincke D A. Intrusion and misuse detection in large-scale systems. IEEE Comput Graph, 2002, 22: 38–47
15 Girardin L. An eye on network intruder-administrator shootouts. In: Proceedings of Workshop on Intrusion Detection and Network Monitoring, Santa Clara, 1999. 19–28
16 Nyarko K, Capers T, Scott C, et al. Network intrusion visualization with niva, an intrusion detection visual analyzer with haptic integration. In: Proceedings of 10th Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems, Orlando, 2002. 277–284
17 Maltego. Paterva Company. http://www.paterva.com/web7
18 Wong T, Jacobson V, Alaettinoglu C. Internet routing anomaly detection and visualization. In: Proceedings of International Conference on Dependable Systems and Networks, Yokohama, 2005. 172–181
19 Teoh S T, Zhang K, Tseng S M, et al. Combining visual and automated data mining for near-real-time anomaly detection and analysis in BGP. In: Proceedings of the 2004 ACM Workshop on Visualization and Data Mining for Computer Security, Washington, 2004. 35–44
20 Teoh S T, Ranjan S, Nucci A, et al. BGP eye: a new visualization tool for real-time detection and analysis of BGP anomalies. In: Proceedings of the 3rd International Workshop on Visualization for Computer Security, Alexandria, 2006. 81–90
21 Arendt D L, Burtner R, Best D M, et al. Ocelot: user-centered design of a decision support visualization for network quarantine. In: Proceedings of IEEE Symposium on Visualization for Cyber Security, Chicago, 2015. 1–8
22 Takada T, Koike H. Tudumi: information visualization system for monitoring and auditing computer logs. In: Proceedings fo 6th International Conference on Information Visualisation, London, 2002. 570–576
23 Ren P, Kristoff J, Gooch B. Visualizing DNS traffic. In: Proceedings of the 3rd International Workshop on Visualization for Computer Security, Alexandria, 2006. 23–30
24 Goodall J R, Lutters W G, Rheingans P, et al. Preserving the big picture: visual network traffic analysis with TN. In: Proceedings of IEEE Workshop on Visualization for Computer Security, Minneapolis, 2005. 47–54
25 Yin X X, Yurcik W, Treaster M, et al. Visflowconnect: netflow visualizations of link relationships for security situational awareness. In: Proceedings of ACM Workshop on Visualization and Data Mining for Computer Security, Washington, 2004. 26–34
26 Zhou F, Huang W, Zhao Y, et al. Entvis: a visual analytic tool for entropy-based network traffic anomaly detection. IEEE Comput Graph Appl, 2015, 35: 1
27 Onut L V, Ghorbani A A. Svision: a novel visual network-anomaly identification technique. Comput Secur, 2007, 26: 201–212
28 Ball R, Fink G A, North C. Home-centric visualization of network traffic for security administration. In: Proceedings of ACM Workshop on Visualization and Data Mining for Computer Security, Washington, 2004. 55–64
29 Lakkaraju K, Yurcik W, Lee A J. Nvisionip: netflow visualizations of system state for security situational awareness. In: Proceedings of ACM Workshop on Visualization and Data Mining for Computer Security, Washington, 2004. 65–72
30 Keim D A, Mansmann F, Schneidewind J, et al. Monitoring network traffic with radial traffic analyzer. In: Proceedings of IEEE Symposium on Visual Analytics Science and Technology, Baltimore, 2006. 123–128
31 Hao L H, Healey C G, Hutchinson S E. Ensemble visualization for cyber situation awareness of network security data. In: Proceedings of IEEE Symposium on Visualization for Cyber Security, Chicago, 2015. 1–8
32 Liao Q, Shi L H, Wang C Y. Visual analysis of large-scale network anomalies. IBM J Res Devel, 2013, 57: 1–12
33 Fink G A, Muessig P, North C. Visual correlation of host processes and network traffic. In: Proceedings of IEEE Workshop on Visualization for Computer Security, Minneapolis, 2005. 11–19
34 Ren P, Gao Y, Li Z, et al. Idgraphs: intrusion detection and analysis using histographs. In: Proceedings of IEEE Workshop on Visualization for Computer Security, Minneapolis, 2005. 39–46
35 McPherson J, Ma K L, Krystosk P, et al. Portvis: a tool for port-based detection of security events. In: Proceedings of ACM Workshop on Visualization and Data Mining for Computer Security, Washington, 2004. 73–81
36 Abdullah K, Lee C, Conti G, et al. Visualizing network data for intrusion detection. In: Proceedings of Information Assurance Workshop From the 6th Annual IEEE SMC, College Park, 2005. 100–108
37 Taylor T, Paterson D, Glanfield J, et al. Flovis: flow visualization system. In: Proceedings of the Cybersecurity Applications & Technology Conference for Homeland Security, Washington, 2009. 186–198

38 Glanfield J, Brooks S, Taylor T, et al. Over flow: an overview visualization for network analysis. In: Proceedings of the International Workshop on Visualization for Cyber Security, Atlantic, 2009. 11–19

39 Zhao Y, Liang X, Fan X P, et al. Mvsec: multi-perspective and deductive visual analytics on heterogeneous network security data. J Visual, 2014, 17: 181–196

40 Fischer F, Mansmann F, Keim D A, et al. Large-scale network monitoring for visual analysis of attacks. In: Proceedings of the 5th International Workshop on Visualization for Computer Security, Cambridge, 2008. 111–118

41 Cortese P F, Battista G D, Moneta A, et al. Topographic visualization of prefix propagation in the internet. IEEE Trans Vis Comput Graph, 2006, 12: 725–732

42 Mansmann F, Daniel A K, Stephen C N, et al. Visual analysis of network traffic for resource planning, interactive monitoring, and interpretation of security threats. IEEE Trans Vis Comput Graph, 2007, 13: 1105–1112

43 Inoue D, Eto M, Suzuki K, et al. Daedalus-viz: novel real-time 3D visualization for darknet monitoring-based alert system. In: Proceedings of the 9th International Symposium on Visualization for Cyber Security, Seattle, 2012. 72–79

44 Inoue D, Eto M, Yoshioka K, et al. Nicter: an incident analysis system toward binding network monitoring with malware analysis. In: Proceedings of WOMBAT Workshop on Information Security Threats Data Collection and Sharing, Amsterdam, 2008. 58–66

45 Oberheide J, Goff M, Karir M. Flamingo: visualizing internet traffic. In: Proceedings of Network Operations and Management Symposium, Vancouver, 2006. 150–161

46 Yelizarov A, Gamayunov D. Visualization of complex attacks and state of attacked network. In: Proceedings of VizSec International Workshop on Visualization for Cyber Security, Atlantic, 2009. 1–9

47 Angelini M, Prigent N, Santucci G. Percival: proactive and reactive attack and response assessment for cyber incidents using visual analytics. In: Proceedings of IEEE Symposium on Visualization for Cyber Security, Chicago, 2015. 1–8

48 Kolaczyk E D, Csrdi G. Visualizing network data. IEEE Trans Vis Comput Graph, 1995, 1: 16–28

49 Matuszak W J, DiPippo L, Sun Y L. Cybersave: situational awareness visualization for cyber security of smart grid systems. In: Proceedings of the 10th Workshop on Visualization for Cyber Security, Atlanta, 2013. 25–32

50 Kotenko I, Novikova E. Visualization of security metrics for cyber situation awareness. In: Proceedings of International Conference on Availability, Reliability and Security, Switzerland, 2014. 506–513

51 Zhao Y, Fan X P, Zhou F F, et al. A survey on network security data visualization. J Comput Aided Des Comput Graph, 2014, 26: 687–697

52 Zhuo W, Nadjin Y. Malwarevis: entity-based visualization of malware network traces. In: Proceedings of the 9th International Symposium on Visualization for Cyber Security, Seattle, 2012. 41–47

53 Trinius P, Holz T, Göbel J, et al. Visual analysis of malware behavior using treemaps and thread graphs. In: Proceedings of 6th International Workshop on Visualization for Cyber Security, Atlantic, 2009. 33–38

54 Gove R, Saxe J, Gold S, et al. Seem: a scalable visualization for comparing multiple large sets of attributes for malware analysis. In: Proceedings of the Eleventh Workshop on Visualization for Cyber Security, Paris, 2014. 72–79

55 Erbacher R F, Christensen K, Sundberg A. Designing visualization capabilities for IDS challenges. In: Proceedings of IEEE Workshop on Visualization for Computer Security, Minneapolis, 2005. 121–127

56 Card S K, Mackinlay J D, Shneiderman B. Readings in Information Visualization: Using Vision to Think. San Francisco: Morgan Kaufmann, 1999

57 Aigner W, Miksch S, Muller W, et al. Visual methods for analyzing time-oriented data. IEEE Trans Vis Comput Graph, 2008, 14: 47–60

58 Xie C, Chen W, Huang X X, et al. VAET: a visual analytics approach for E-transactions time-series. IEEE Trans Vis Comput Graph, 2014, 20: 1743–1752

59 Kondo B, Collins C M. Dimpvis: exploring time-varying information visualizations by direct manipulation. IEEE Trans Vis Comput Graph, 2014, 20: 2003–2012

60 Isaacs K E, Bremer P T, Jusufi I, et al. Combing the communication hairball: visualizing parallel execution traces using logical time. IEEE Trans Vis Comput Graph, 2014, 20: 2349–2358

61 Gotz D, Stavropoulos H. Decisionflow: visual analytics for high-dimensional temporal event sequence data. IEEE Trans Vis Comput Graph, 2014, 20: 1783–1792

62 Cho I, Dou W, Wang D X Y, et al. Vairoma: a visual analytics system for making sense of places, times, and events in Roman history. IEEE Trans Vis Comput Graph, 2016, 22: 210–219

63 Fulda J, Brehmer M, Munzner T. Timelinecurator: interactive authoring of visual timelines from unstructured text. IEEE Trans Vis Computer Graph, 2016, 22: 300–309

64 Loorak M H, Perin C, Kamal N, et al. Timespan: using visualization to explore temporal multi-dimensional data of stroke patients. IEEE Trans Vis Comput Graph, 2016, 22: 409–418

65 Walker J, Borgo R, Jones M W. Timenotes: a study on effective chart visualization and interaction techniques for time-series data. IEEE Trans Vis Comput Graph, 2016, 22: 549–558

66 Bach B, Shi C, Heulot N, et al. Time curves: folding time to visualize patterns of temporal evolution in data. IEEE Trans Vis Comput Graph, 2016, 22: 559–568

67 Gu Y, Wang C L, Peterka T, et al. Mining graphs for understanding time-varying volumetric data. IEEE Trans Vis Comput Graph, 2016, 22: 965–974

68 Albo Y, Lanir J, Bak P, et al. Off the radar: comparative evaluation of radial visualization solutions for composite indicators. IEEE Trans Vis Comput Graph, 2016, 22: 569–578

69 Gschwandtner T, Bogl M, Federico P, et al. Visual encodings of temporal uncertainty: a comparative user study.

IEEE Trans Vis Comput Graph, 2016, 22: 539–548
70 Sun G D, Wu Y C, Liu S X, et al. Evoriver: visual analysis of topic coopetition on social media. IEEE Trans Vis Comput Graph, 2014, 20: 1753–1762
71 Heimerl F, Han Q, Koch S. Citerivers: visual analytics of citation patterns. IEEE Trans Vis Comput Graph, 2016, 22: 190–199
72 Zhao J, Cao N, Wen Z, et al. Fluxflow: visual analysis of anomalous information spreading on social media. IEEE Trans Vis Comput Graph, 2014, 20: 1773–1782
73 Chen W, Guo F Z, Wang F Y. A survey of traffic data visualization. IEEE Trans Intel Transp Syst, 2015, 16: 2970–2984
74 Gratzl S, Gehlenborg N, Lex A, et al. Domino: extracting, comparing, and manipulating subsets across multiple tabular datasets. IEEE Trans Vis Comput Graph, 2014, 20: 2023–2032
75 Kim H, Choo J, Park H, et al. Interaxis: steering scatterplot axes via observation-level interaction. IEEE Trans Vis Comput Graph, 2016, 22: 131–140
76 Lowe T, Forster E C, Albuquerque G, et al. Visual analytics for development and evaluation of order selection criteria for autoregressive processes. IEEE Trans Vis Comput Graph, 2016, 22: 151–159
77 Chen W, Shen Z Q, Tao Y B. Data Visualization. Beijing: Publishing House of Electronic Industry, 2013
78 Cao N, Shi C, Lin S, et al. Targetvue: visual analysis of anomalous user behaviors in online communication systems. IEEE Trans Vis Comput Graph, 2016, 22: 280–289
79 Rubio-Sanchez M, Raya L, Diaz F, et al. A comparative study between radviz and star coordinates. IEEE Trans Vis Comput Graph, 2016, 22: 619–628
80 Papadopoulos C, Gutenko I, Kaufman A E. Veevvie: visual explorer for empirical visualization, vr and interaction experiments. IEEE Trans Vis Comput Graph, 2016, 22: 111–120
81 Wang J, Mueller K. The visual causality analyst: an interactive interface for causal reasoning. IEEE Trans Vis Comput Graph, 2016, 22: 230–239
82 Lee S, Kim S H, Hung Y H, et al. How do people make sense of unfamiliar visualizations?: a grounded model of novice's information visualization sensemaking. IEEE Trans Vis Comput Graph, 2016, 22: 499–508
83 Johansson J, Forsell C. Evaluation of parallel coordinates: overview, categorization and guidelines for future research. IEEE Trans Vis Comput Graph, 2016, 22: 579–588
84 Raidou R G, Eisemann M, Breeuwer M, et al. Orientation-enhanced parallel coordinate plots. IEEE Trans Vis Comput Graph, 2016, 22: 589–598
85 Chen H D, Zhang S, Chen W, et al. Uncertainty-aware multidimensional ensemble data visualization and exploration. IEEE Trans Vis Comput Graph, 2015, 21: 1072–1086
86 Roberts J C, Headleand C, Ritsos P D. Sketching designs using the five design-sheet methodology. IEEE Trans Vis Comput Graph, 2016, 22: 419–428
87 VanderPlas S, Hofmann H. Spatial reasoning and data displays. IEEE Trans Vis Comput Graph, 2016, 22: 459–468
88 Goodwin S, Dykes J, Slingsby A, et al. Visualizing multiple variables across scale and geography. IEEE Trans Vis Comput Graph, 2016, 22: 599–608
89 Scheepens R, Hurter C, van de Wetering H, et al. Visualization, selection, and analysis of traffic flows. IEEE Trans Vis Comput Graph, 2016, 22: 379–388
90 Lehmann D J, Theisel H. Optimal sets of projections of high-dimensional data. IEEE Trans Vis Comput Graph, 2016, 22: 609–618
91 Cheng S H, Mueller K. The data context map: fusing data and attributes into a unified display. IEEE Trans Vis Comput Graph, 2016, 22: 121–130
92 Jackle D, Fischer F, Schreck T, et al. Temporal mds plots for analysis of multivariate data. IEEE Trans Vis Comput Graph, 2016, 22: 141–150
93 Stahnke J, Dork M, Muller B, et al. Probing projections: interaction techniques for interpreting arrangements and errors of dimensionality reductions. IEEE Trans Vis Comput Graph, 2016, 22: 629–638
94 Kohonen T. Self-Organizing Maps. New York: Springer, 1997. 266–270
95 Amini F, Rufiange S, Hossain Z, et al. The impact of interactivity on comprehending 2D and 3D visualizations of movement data. IEEE Trans Vis Comput Graph, 2015, 21: 122–135
96 Tory M, Kirkpatrick A E, Atkins M S, et al. Visualization task performance with 2D, 3D, and combination displays. IEEE Trans Vis Comput Graph, 2006, 12: 2–13
97 Sun M Y, Mi P, North C, Ramakrishnan N. Biset: semantic edge bundling with biclusters for sensemaking. IEEE Trans Vis Comput Graph, 2016, 22: 310–319
98 Von Landesberger T, Brodkorb F, Roskosch P, et al. Mobilitygraphs: visual analysis of mass mobility dynamics via spatio-temporal graphs and clustering. IEEE Trans Vis Comput Graph, 2016, 22: 11–20
99 Krause J, Perer A, Bertini E. Infuse: interactive feature selection for predictive modeling of high dimensional data. IEEE Trans Vis Comput Graph, 2014, 20: 1614–1623
100 Mahyar N, Tory M. Supporting communication and coordination in collaborative sensemaking. IEEE Trans Vis Comput Graph, 2014, 20: 1633–1642
101 Stolper C D, Perer A, Gotz D. Progressive visual analytics: user-driven visual exploration of in-progress analytics. IEEE Trans Vis Comput Graph, 2014, 20: 1653–1662
102 Klemm P, Oeltze-Jafra S, Lawonn K, et al. Interactive visual analysis of image-centric cohort study data. IEEE Trans Vis Comput Graph, 2014, 20: 1673–1682

103 Jang S, Elmqvist N, Ramani K. Motionflow: visual abstraction and aggregation of sequential patterns in human motion tracking data. IEEE Trans Vis Comput Graph, 2016, 22: 21–30
104 Nguyen P H, Xu K, Wheat A, et al. Sensepath: understanding the sensemaking process through analytic provenance. IEEE Trans Vis Comput Graph, 2016, 22: 41–50
105 Blascheck T, John M, Kurzhals K, et al. Va2: a visual analytics approach for evaluating visual analytics applications. IEEE Trans Vis Comput Graph, 2016, 22: 61–70
106 Kwon B C, Kim S H, Lee S, et al. Visohc: designing visual analytics for online health communities. IEEE Trans Vis Comput Graph, 2016, 22: 71–80
107 Glueck M, Hamilton P, Chevalier F, et al. Phenoblocks: phenotype comparison visualizations. IEEE Trans Vis Comput Graph, 2016, 22: 101–110
108 Guo H, Gomez S R, Ziemkiewicz C, et al. A case study using visualization interaction logs and insight metrics to understand how analysts arrive at insights. IEEE Trans Vis Comput Graph, 2016, 22: 51–60