社団法人　情報処理学会　研究報告
IPSJ SIG Technical Reports

2006－CSEC－34 (63)
2006／7／21

テクスチャキューブを用いた対話的ネットワークトラヒック視覚化ツール

Erwan Le Malécot[†]　小原　正芳[††]　堀　良彰[††]　櫻井　幸一[††]

[†] 九州システム情報技術研究所, 814-0001 福岡市早良区百道浜 2-1-22 — Supélec, Avenue de la Boulaie, BP81127, 35511 Cesson-Sévigné Cedex, France
[††] 812-8581 福岡市東区箱崎 6-10-1 九州大学大学院システム情報科学府
E-mail: [†]lemalecot@isit.or.jp,erwan.lemalecot@supelec.fr,
[††]kohara@itslab.csce.kyushu-u.ac.jp,
{hori,sakurai}@csce.kyushu-u.ac.jp

あらまし　トラヒックモニタリングと解析を効率的に行える3次元表現を用いたネットワークトラヒック視覚化ツールを設計した。本ツールは、視覚化対象となるネットワークをテクスチャキューブを用いて立体的に表現する。ネットワークにおける通信は、キューブ空間中のテクスチャおよびキューブ間のリンクによってトラヒックの属性を色分けし表現する。さらに、本ツールのプロトタイプ実装について述べ、その評価を行う。

キーワード　視覚化, 3D, トラヒックモニタリング, ネットワークセキュリティ

Interactive Textured Cube Based Network Traffic Visualization for Network Monitoring and Security

Erwan LE MALÉCOT[†], Masayoshi KOHARA[††], Yoshiaki HORI[††], and Kouichi SAKURAI[††]

[†] Institute of Systems & Information Technologies/KYUSHU, 2-1-22, Momochihama, Sawara-ku, Fukuoka, 814-0001, Japan — Supélec, Avenue de la Boulaie, BP81127, 35511 Cesson-Sévigné Cedex, France
[††] Graduate School of Information Science and Electrical Engineering, Kyushu University, 6-10-1 Hakozaki, Higashi-ku, Fukuoka, 812-8581 Japan
E-mail: [†]lemalecot@isit.or.jp,erwan.lemalecot@supelec.fr,
[††]kohara@itslab.csce.kyushu-u.ac.jp,
{hori,sakurai}@csce.kyushu-u.ac.jp

Abstract We introduce an original visualization design based on an interactive 3D representation of the traffic between selected computer networks in order to achieve the efficient monitoring and analysis of this traffic. This design is based on a representation of the selected networks by textured cubes. Textures are modified according to the activity of the associated networks. Network traffic is then displayed as links between the cubes, the geometrical and color properties of these links being used to encode various properties of the traffic. We also present a prototype implemented according to this design and tests that were made to evaluate it.

Key words Visualization, 3D, Network, Monitoring, Security.

1. Introduction

Nowadays, many organizations heavily rely on computer networks which then become critical parts of their infrastructures. At the same time, the number of malicious attacks against such networks is exploding. In order to guarantee the security and the stability of the networks they manage, system administrators must monitor them efficiently. That monitoring task can be done displaying selected information about the network traffic in a textual form and reading this output directly. Unfortunately, the amount of data captured on the network they need to analyze to perform that task tends to become quite huge and that method is no longer really feasible. Various tools exist to help system administrators dealing with this issue. They are mainly based on automated systems (learning techniques, signature databases or statistical analysis).

Recently, visualization techniques also started to be used for the monitoring task. The idea is to take advantage of human visual processing and pattern recognition that are rather unexploited resources in traditional tools. We believe that the use of visualization techniques can greatly improve the detection of network attacks and anomalies. Various visualization tools for network monitoring already exist. Most of these tools are based on 2D visualization display zones. Some try to exploit 3D, however they generally face clarity and usability issues.

In this paper, we introduce an original visualization design for network monitoring. It is based on an interactive 3D representation of the network traffic between selected network zones. These zones are represented by textured cubes. The textures are modified according to the activity of the associated network zones. The network traffic is then displayed as links between the cubes, the geometrical and color properties of these links being used to encode various properties of this traffic. As a 3D representation is sometimes difficult to understand and analyze, we chose to complete it by an additional 2D representation. That 2D representation corresponds to a projection of some selected parts of the 3D one. It is based on an original interactive grid system to represent the network space. A prototype, which is also introduced in this paper, was made in order to test the proposed visualization design.

This paper is organized as follows. In Section 2, we introduce some works done in the field of information visualization. We also describe and analyze selected existing visualization designs for network traffic monitoring and network security. In Section 3, we explain our original visualization design. In Section 4, we introduce the prototype we built following this visualization design and present some interesting patterns we were able to visualize with it. In Section 5, we conclude on the use of that prototype and finally in Section 6, we discuss some future work we plan to do in relation with the proposed visualization design.

2. Related work

The idea of applying visualization techniques to the field of network security and more especially to network traffic monitoring is quite recent. However, research has been active in the field of information visualization for a lot longer. Shneiderman proposes a taxonomy of information visualization techniques sorting them according to the type of data visualized but also according to the task to be accomplished [14]. It provides a large panel of available basic techniques that can be combined to make complex visualization systems. He also introduces the mantra: "Overview first, zoom and filter, then details-on-demand". We believe that this mantra provides a good summary of the needed functionalities for a visualization system. Chi proposes a similar taxonomy but more complete and recent. It is also based on different criteria [8].

Several tools already exist to monitor networks graphically. Most of them are based on a 2D visualization design. For instance, EtherApe [2] is a tool that graphically displays network activity between hosts. It displays hosts as nodes laid out in a circle. Network activity is symbolized by the thickness of the links between the nodes and colors are used to represent the recognized protocols. This tool is well-suited for a limited number of hosts and connections. But as the number of hosts increases, the visualization zone becomes difficult to interpret. Also, the nodes are reorganized during the display of network activity and that makes it complex to locate similar patterns of network activity between several captures.

VISUAL [7] is another 2D visualization system. System administrators are interested in the interactions between the network they manage and the outside world, Ball et al. chose to focus on this monitored network for their visualization design. The monitored network is displayed as a grid whose quadrilaterals each represent a host. External hosts are represented as markers laid around the grid. Connections between internal and external hosts are symbolized by segments linking the markers and the associated quadrilaterals. The color of the segment indicates if the communication is bidirectional or unidirectional (in this case the direction is also indicated). The communication bandwidth is encoded in the size of the markers. VISUAL also proposes filtering functionalities and it is possible to retrieve detailed information about a host by selecting it in the visualization zone. VISUAL provides an efficient view of the interactions between the monitored network and the external network. However a lot of network attacks also come from inside the monitored network and so, its internal activity should also be displayed.

As for 3D visualization designs, the Spinning Cube of Potential Doom [11] displays data (provided by the Bro IDS [1]) regarding established TCP connections and connection attempts on a cube. The x-axis of this cube stands for the local network's IP address space, the z-axis for the global IP address space and the y-axis for the TCP ports. Each TCP connection is displayed by a dot. This dot is white for an established connection and colored according to the port number for a connection attempt. In this way, connection attempts are emphasized as they are usually the mark of port scans. This visualization system successfully uses a 3D representation that permits easy detection of a specific kind of attack. However, it does not enable the user to fully interact with the visualization zone, for instance to zoom and find the origin of an attack.

IDtk [10] is another 3D visualization system. It can process either Snort IDS [6] output (alerts) or raw TCP network traffic. IDtk is based on the display of glyphs whose attributes are used to represent the components of the input data. These components can be encoded in the coordinates, the color, the size and the shape of the glyphs. So with this tool, the user is free to customize the visualization zone. Despite that, the design is limited as it only permits the detection of relationships between elements of the data (attributes of the corresponding glyphs will be close or far).

Oline and Reiners [13] also propose several 3D visualization designs for network security using either IDS output or raw network traffic (TCP and UDP).

Finally, as we are interested in visualization systems for network security, we believe that it is important to keep in mind that the visualization systems themselves can be the targets of attacks. Conti et al. [9] deal with this issue and provides an useful catalog of possible attacks against visualization systems.

3. The visualization design

The current visualization design derives from a 2D visualization design we previously proposed [12]. It was built to overcome some usability and scalability issues that we encountered with the previous design. This previous design was built to provide a system administrator with both a view of the internal traffic of the network he monitors and a view of the interactions of this network with the outside world. The visualization design we propose in this paper has the same purpose.

It is based on two different representations of the network traffic, one in 2D and the other in 3D. Both of them are based on an interactive grid representation of selected network zones. The 2D representation consists of two of these grids placed beside. Hosts are mapped on the grids and the network traffic is displayed as lines joining the parts of the grids standing for hosts currently communicating. The grids are also colored according to the amount of activity of the mapped hosts. It is basically the same for the 3D representation except that the grids are mapped as textures on the faces of several cubes.

The design we propose in this paper is adapted to the IPv4 network address space. The information that we display on this design is to be extracted from IP network packets. It includes: the source IP address, the destination IP address, the source port, the destination port and the protocol. We are also displaying the time of capture of the IP packets as it is an essential piece of information for network traffic monitoring and analysis.

3.1 The grid representation

The core element of our visualization design is a hierarchical representation of the network space based on interactive 16x16 grids (Fig. 1). Using this system, the user can access and visualize the activity of any part of the global network.

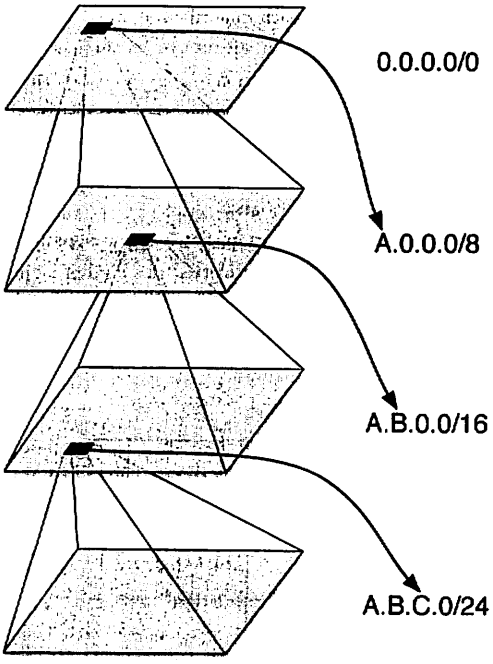


Fig. 1 Interactive grid system.

Hosts are mapped on the grids according to their IP addresses. In the IPv4 system, IP addresses are coded by 32 bits. An usual way to represent them is to use the dotted quad notation, IP addresses are written in the form X.X.X.X where each X stands for a byte (8 bits). The initial state of the grids is a mapping of the IP addresses on these grids based on their first byte. It means that one quadrilateral of the grids is associated with all the hosts whose IP addresses share the same first byte. Then, from this state, any /24 network can be accessed by gradually selecting the bytes of the network address. For instance, if a user selects the quadrilateral standing for the value 192 on a grid, a new grid is displayed, replacing the previous one, that represents the network zone 192.0.0.0/8. Each quadrilateral of this new grid is associated with the set of IP addresses whose first byte equals 192 and whose second byte equals the value mapped on the quadrilateral. In the same way, the user can select the second byte, for example 168, the grid is refreshed to represent the 192.168.0.0/16 zone and he can finally select the third byte. At this point, each quadrilateral of the grid stands for an unique host.

So in the initial state, the grids provide an overview of the global network activity. The data is quite aggregated and each quadrilateral stands for a lot of IP addresses. From this state, the user can modify the interactive grids to make them correspond to smaller network zones and so access more detailed information as each quadrilateral stand for less IP addresses. Also, one of the advantages of our hierarchical representation is that the position of a given host on the grid system does not vary. If several data sets are displayed on a visualization system using this kind of representation, the detection of specific visual patterns is easier as the network space representation is constant.

As for the mapping algorithm, bytes are mapped using a quadtree scheme (Fig. 2). We selected this kind of mapping because it has an interesting property: if it is used to map IP addresses, it gathers them in subnets on the grids. In a way, it permits to represent the logical network topology, which is quite interesting for system administrators.

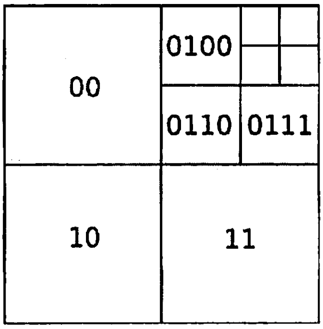


Fig. 2 Quadtree mapping.

3.2 The 2D representation

The 2D representation is made of two interactive grids which behave as described in Section 3.1. One of the grids stands for the hosts as sources (referred to as source grid from now on) and the other one for the hosts as destinations (referred as destination grid from now on). The network traffic is then displayed on that network representation as lines joining the source and destination zones. One line is drawn per packet captured and the lines are colored according to the protocol of the packets. As for network activity, the quadrilaterals of the grids are colored according to the number of packets sent (for the source grid) or received (for the destination grid) by the associated network zones. The higher the number of packets, the more vivid the color.

3.3 The 3D representation

The 3D representation is based on textured cubes laid out in a 3D space. Each cube stands for a network zone which can be either a /0, /8, /16 or a /24 network zone. The faces of the cubes are textured with grids. Like the one used for the 2D representation, these grids behave as presented in Section 3.1. For instance, if a cube stands for a /8 network, each quadrilateral of the mapped grids represents a /16 subnet. Three faces of each cube are mapped with the source grid corresponding to the network the cube represents (referred to as source faces from now on) and the other three faces are mapped with the destination grid (referred to as destination faces from now on).

The activity of the different network zones are represented the same way that for the grids of the 2D representation, encoded in the color of the quadrilaterals of the mapped grids. Two different color scales are used in order to differentiate the source faces and the destination faces.

A source face of a cube can be connected to a destination face of another cube. In this case, the traffic between the network zone represented by the first cube and the one represented by the second cube is displayed on the visualization zone. For each packet captured, a is drawn between the quadrilateral standing for the source network zone (on the texture of the selected source face) and the quadrilateral standing for the destination network zone (on the texture of the selected destination face). The use of cones permits to simply visualize the direction of propagation of the packets. A transparency effect (Fig. 6) is used to display the time of capture information associated with each packet. The older the packet, the more transparent the cone. In addition to that, the color scale used to indicate the protocol of a captured packet in the 2D representation is also used for the color of the associated cone.

3.4 Interactions between the 2D and the 3D representations

The captured network traffic is displayed simultaneously on both 2D and 3D representations. The 3D representation can provide an overview of the traffic between several selected network zones as several cubes can be placed on the 3D space. Regarding the 2D representation, it can provide a detailed view of the traffic between two selected network zones. So the idea is to combine efficiently the two proposed representations as they provide two complementary functionalities.

A pair of two cubes can be selected on the 3D representation and the traffic between the two associated network zones is then displayed on the 2D representation. Similarly, a pair of two cubes of the 3D representation can be synchronized with the grids positions of the 2D representation so that the network zones represented by the grids also become the network zones represented by the cubes. It enables the user of the visualization system to analyze finely (using the

2D representation) suspicious network traffic he detects on the overview (3D representation). If this suspicious network traffic happens to be malicious or quite unusual, the user can modify the settings of the overview in order to finely monitor that suspicious traffic directly from it.

4. The prototype

4.1 Implementation

In order to test the accuracy of our approach, we implemented a prototype written in C language that is based the proposed visualization design. It uses the GTK+-2.X [3] library and GtkGLArea [4] (OpenGL widget for GTK+-2.X) for the graphic user interface (GUI), and the Libpcap [5] library for the packet capture. It is known to compile and run on FreeBSD, GNU/Linux and Mac OS X operating systems.

4.2 User interface

In order to separate various functionalities of the prototype, we chose to organize the interface in four tabs:

- The "Visualization" tab (Fig. 3): This is the main screen of the prototype. From this part, the user can visualize both the 2D (at the right side) and the 3D (at the left side) representations of the network traffic. At the center, the user can access several controls which are also grouped in several tabs:
  - The "C" (standing for "Capture Controls") tab: It provides the controls to start a live packet capture and to open a previously saved network traffic dump file.
  - The "V" (standing for "Visualization Controls") tab: It provides some controls regarding the two visualization zones (reset them to their original states, refresh the display, ...).
  - The "G" (standing for "Grid & Cube Controls") tab: It provides controls to select cubes on the 3D visualization zone, to synchronize the 2D and the 3D visualization zones, ...
  - The "S" (standing for "Scale") tab: It provides the color scale (association between colors and protocols) used for the cones and the lines representing the packets.
- The "Messages" tab: Some status messages are printed under this tab. During the packet capture, packet headers are also logged here so that the user can access this kind of information if he needs to.
- The "Preferences" tab (Fig. 4): The user can set some options and modify the behavior of the program. For instance, he can modify the color scale or set a filter to display only matching packets on the visualization zones.
- The "Help" tab: To display some basic information about the prototype.

4.3 Use of the prototype

For the tests we chose to display five cubes on the 3D visualization zone. However, it can be modified and adapted to the user's needs. In the initial state, the cubes at the left, middle and right all stand for the local network (a /24 network zone). The cubes at the top and at the bottom both stand for the global network (the /0 network zone). We kept that configuration for Fig. 5, Fig. 6 and Fig. 7. So the internal traffic of the local network is displayed between the left and the middle cubes. It is also displayed between the middle and right cubes. The network traffic displayed between the bottom and the middle cubes is the incoming traffic from the global network to the local network. The network traffic displayed between the middle and the top cubes is the outgoing traffic from the local network to the global network.

For Fig. 5, we used a small capture done on our local network from a connected host. The 2D visualization zone is set to display the internal traffic of the local network. The observed traffic is quite common, the local traffic is mainly constituted of broadcast activity. We can also notice some POP3 packets to our mail server. For Fig. 6, the transparency effect is activated and we can see the cones standing for the packets vanishing on the 3D visualization zone.

Regarding Fig. 7, we probed all the local network on the port 22 (SSH) from a local host. The portsweep is highly noticeable on the 3D visualization zone and the 2D visualization zone then enables the user to get the exact source of that attack.

5. Conclusion

This paper introduces a visualization design combining both 2D and 3D representations of the network traffic. This design is inspired by a previous one we proposed which had some scalability and usability issues. A prototype was made based on that improved design in order to test it. We used this prototype not only to display several network traffic dumps but also to visualize network traffic in real time. We were able to detect several patterns corresponding to abnormal activity with the help of the prototype. Using that kind of visualization system really simplifies the detection of such patterns in the data compared to classical textual based techniques. The proposed visualization design has proven to be scalable through the 3D representation of the network traffic. Also, interactively combining both 2D and 3D representations lets the user finely control what he is visualizing. We believe that using our prototype can significantly improve the efficiency of the network traffic monitoring task.

6. Future work

The current prototype includes all the basic functionalities corresponding to the proposed visualization design. However, some further functionalities can be added to simplify

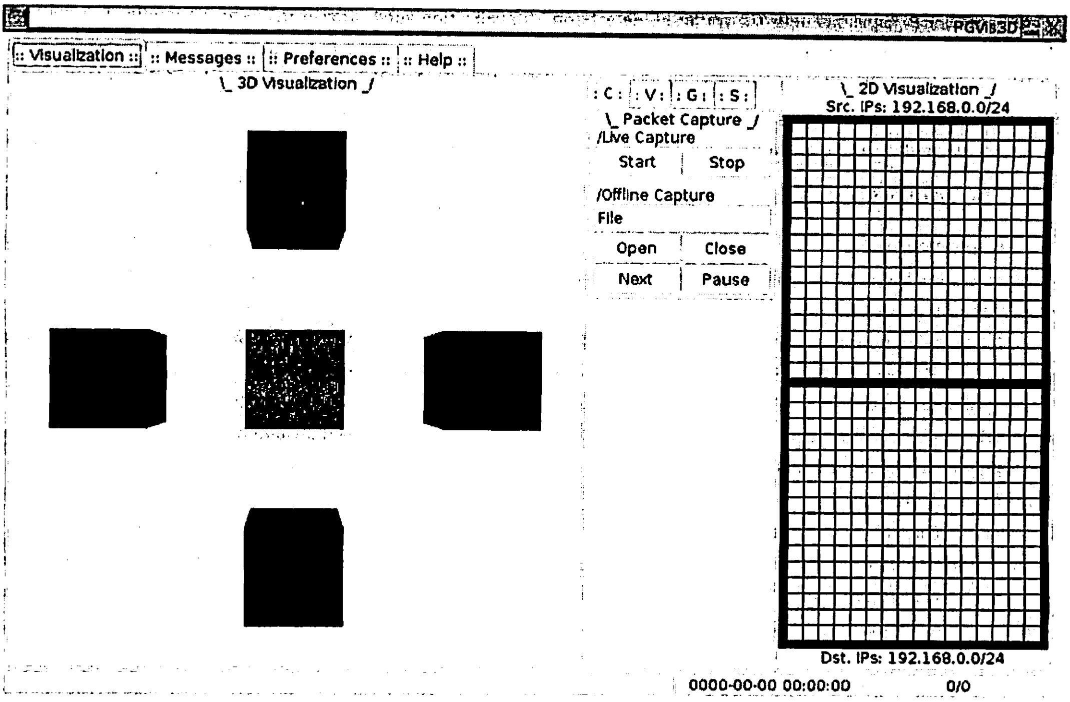
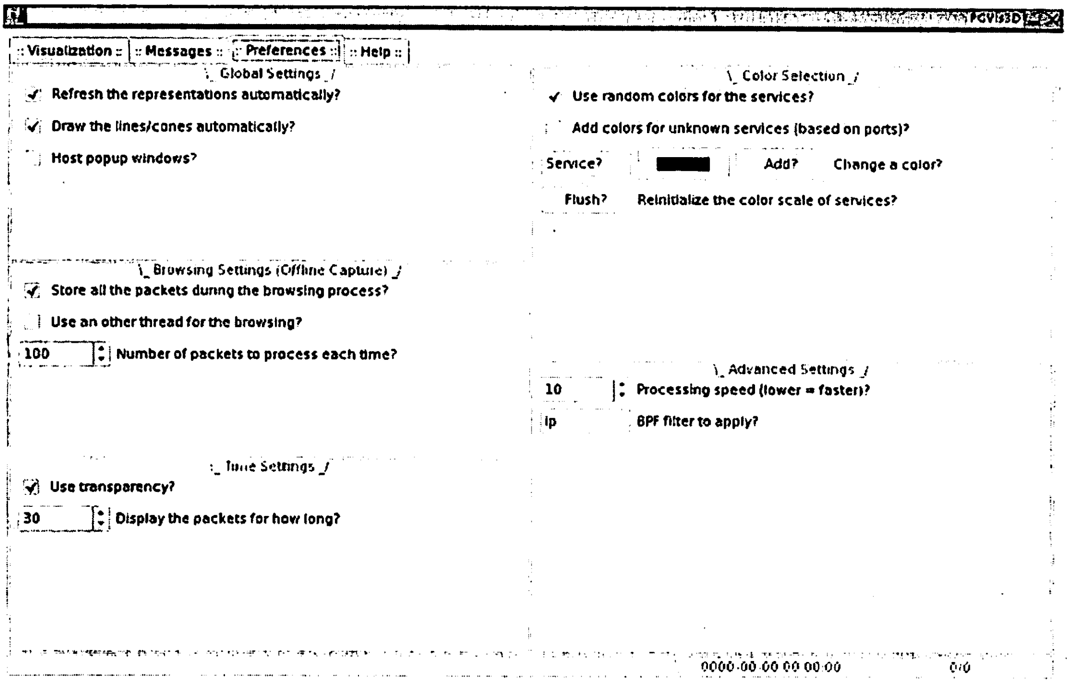


Fig. 3 The main screen of the prototype.



_ Global Settings _/

☑ Refresh the representations automatically?

☑ Draw the lines/cones automatically?

☐ Host popup windows?

_ Browsing Settings (Offline Capture) _/

☑ Store all the packets during the browsing process?

☐ Use an other thread for the browsing?

100 Number of packets to process each time?

_ Time Settings _/

☑ Use transparency?

30 Display the packets for how long?

_ Color Selection _/

✓ Use random colors for the services?

☐ Add colors for unknown services (based on ports)?

Service? ■ Add? Change a color?

Flush? Reinitialize the color scale of services?

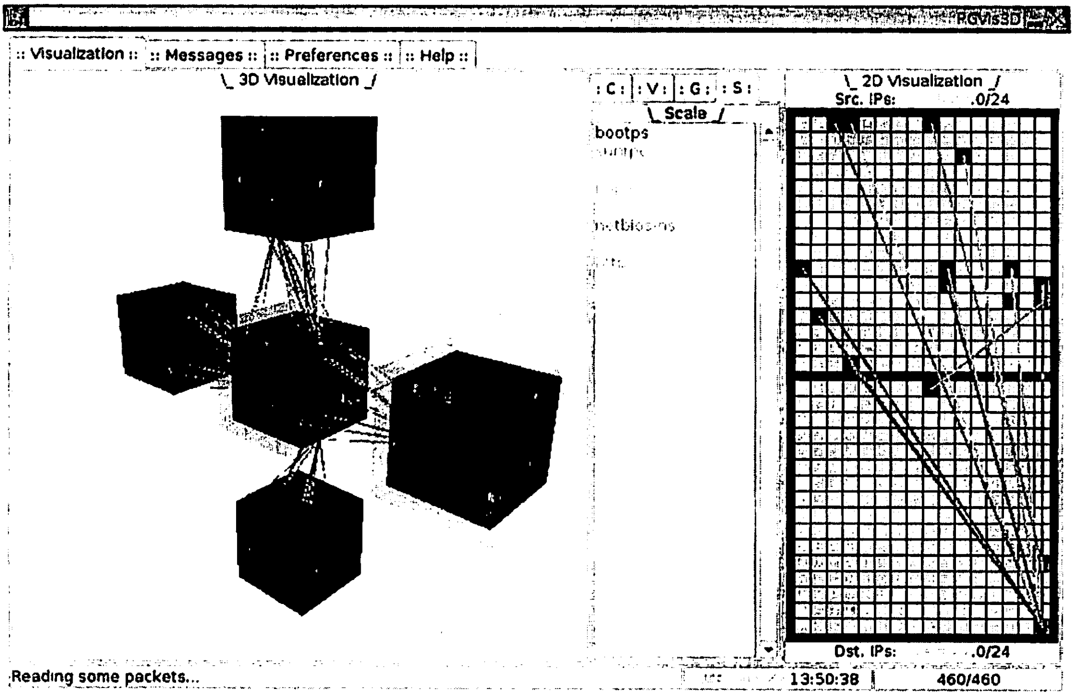
_ Advanced Settings _/

10 Processing speed (lower = faster)?

ip BPF filter to apply?

Fig. 4 The "Preferences" tab.

its use or help the user in the network traffic monitoring task. For instance, we plan to add an history of the network zones that were displayed on the 2D visualization zone in order to provide a quick access to them. We also plan to improve the



Reading some packets...

Fig. 5 Display of some usual network traffic on the prototype.

management of the 3D visualization zone to enable the user to easily modify the disposition and the number of the cubes in order to meet his needs (the current system being quite complex).

An other point we are interested in is the adaptation of our visualization design to the IPv6 address space as it is to replace the IPv4 system.

Acknowledgments

Support for this work is partially provided by National Institute of Information Communications and Technology (NICT), Japan, as a visualization technology research for the *nicter* Incident Analysis Center project.

References

[1] Bro. http://bro-ids.org.

[2] Etherape. http://etherape.sourceforge.net.

[3] Gtk+ library. http://www.gtk.org.

[4] Gtkglarea widget. http://www.student.oulu.fi/~jlof/gtkglarea/.

[5] Libpcap library. http://www.tcpdump.org.

[6] Snort. http://www.snort.org.

[7] Robert Ball, Glenn A. Fink, and Chris North. Home-centric visualization of network traffic for security administration. In *Proceedings of the 2004 ACM workshop on Visualization and data mining for computer security (VizSEC/DMSEC'04)*, pages 55-64, New York, NY, USA, 2004. ACM Press.

[8] Ed H. Chi. A Taxonomy of Visualization Techniques Using the Data State Reference Model. In *Proceedings of the IEEE Symposium on Information Vizualization (INFOVIS '00)*, page 69, Washington, DC, USA, 2000. IEEE Computer Society.

[9] Gregory Conti, Mustaque Ahamad, and John Stasko. Attacking information visualization system usability overloading and deceiving the human. In *Proceedings of the Symposium on Usable Privacy and Security (SOUPS'05)*, pages 89-100, New York, NY, USA, 2005. ACM Press.

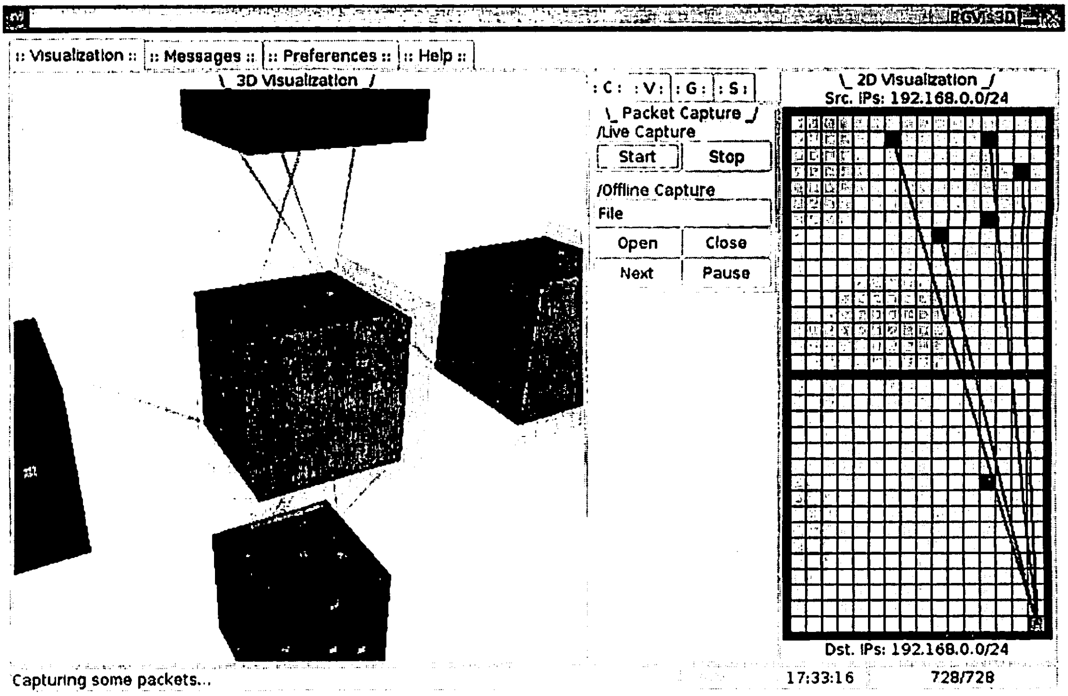
[10] Anita Komlodi, Penny Rheingans, Utkarsha Ayat, John R. Goodall, and Amit Joshi. A User-centered Look at Glyph-based Security Visualization. In *Proceedings of the IEEE Workshops on Visualization for Computer Security (VizSEC'05)*, pages 21-28, Washington, DC, USA, 2005. IEEE Computer Society.

[11] Stephen Lau. The Spinning Cube of Potential Doom. *Commun. ACM*, 47(6):25-26, 2004.

[12] Erwan Le Malécot, Masayoshi Kohara, Yoshiaki Hori, and Kouichi Sakurai. Grid Based Network Address Space Browsing for Network Traffic Visualization. In *Proceedings of the 7th IEEE Information Assurance Workshop*, pages 261-267, Washington, DC, USA, 2006. IEEE Computer Society.

[13] Adam Oline and Dirk Reiners. Exploring Three-dimensional Visualization for Intrusion Detection. In *Proceedings of the IEEE Workshops on Visualization for Computer Security (VizSEC'05)*, pages 113-120, Washington, DC, USA, 2005. IEEE Computer Society.

[14] Ben Shneiderman. The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations. In *Proceedings of the IEEE Symposium on Visual Languages (VL'96)*, pages 336-343, 1996.



Capturing some packets...

Fig. 6 Transparency effect to represent the time of capture information.

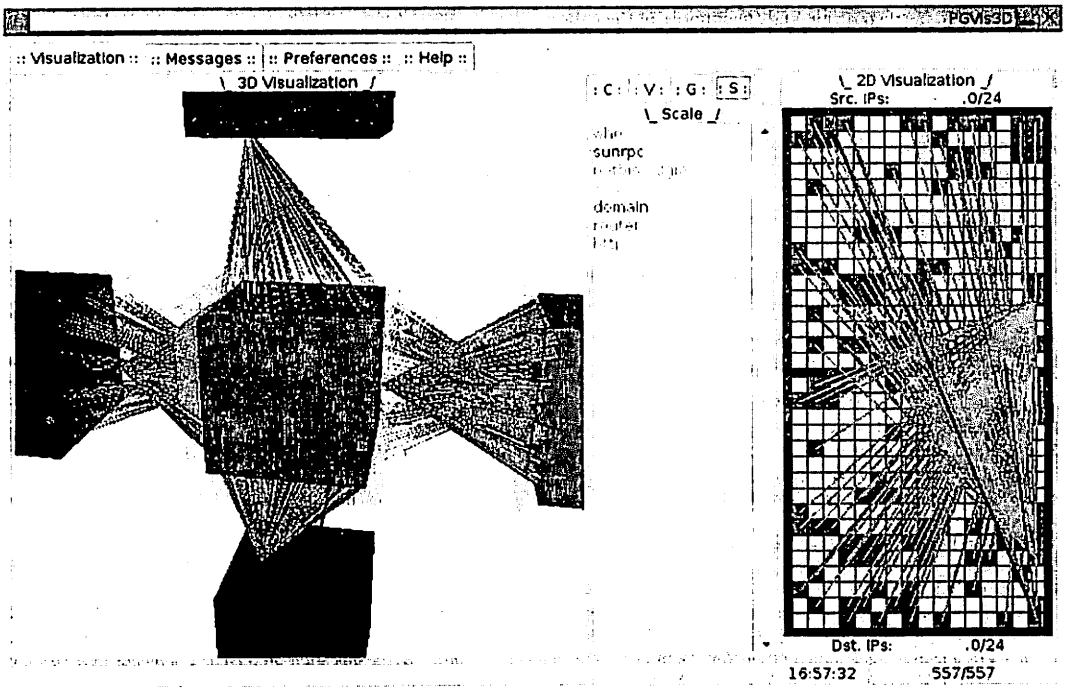


Fig. 7 A network scan from an internal host.