VAST: Visualizing Autonomous System Topology

Jon Oberheide
Networking Research and Development
Merit Network Inc.
Ann Arbor, MI 48105
jonojono@umich.edu

Manish Karir
Networking Research and Development
Merit Network Inc.
Ann Arbor, MI 48105
mkarir@merit.edu

Dionysus Blazakis
Hybrid Networks Center
Institute for Systems Research
University of Maryland
College Park, MD 20742
dblaze@isr.umd.edu

ABSTRACT

Extracting specific and relevant information regarding the Internet's BGP routing topology is a challenging task. In this paper we present a set of techniques that can be used to visualize various aspects of the Internet topology. We have implemented our visualization techniques in a unique tool called VAST (Visualizing Autonomous System Topology). With the help of simple illustrative examples we describe how our visualizations allow security researchers to extract relevant information quickly from raw BGP routing datasets. VAST provides visualizations that represent information about both the overall topological properties of the Internet as well as individual Autonomous System (AS) behavior.

1. INTRODUCTION

Raw BGP routing data is fairly complex and therefore it is difficult to gain a good understanding of the underlying information that it contains. Visualization is a technique that is well suited to this problem as a network operator can often easily apply intuitive understanding and detailed knowledge of the network in order to identify features that are difficult for an automated system to detect.

Recently, the problem of understanding the Internet's BGP AS topology has gained prominence. An understanding of the Autonomous System (AS) topology and inter-AS relationships is important for many reasons such as understanding peering relationships, peering evaluations, and routing hot-spot detection. One particular application of this research is in gaining a topological understanding of the vulnerability of the Internet and the identification of critical Internet infrastructure. Such an understanding can help in the planning for various contingencies in the face of widespread worms, viruses or other malicious attacks.

In this paper we develop visualization techniques that can help us gain a better understanding about the topological properties of the Internet. We have implemented our techniques in a tool called VAST (Visualizing Autonomous System Topology). VAST takes raw BGP routing datasets as input such as the extensive archives at Routeviews [1]. The raw data is processed, relevant statistics are extracted and stored in various tree-based data structures. This processed data is then suitable for visualization.

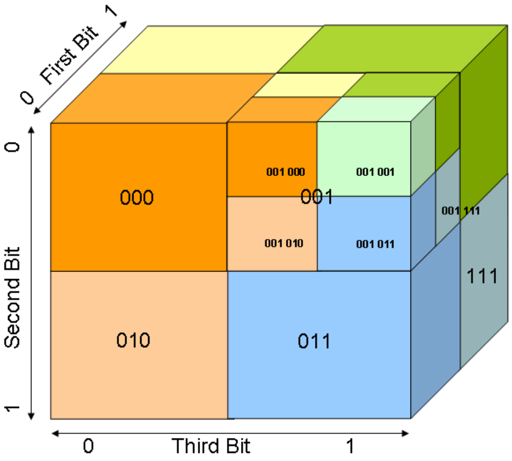


Figure 1: Octo-tree Algorithm: Using 3 bits at a time we can specify one particular sub-cube of a larger cube. Repeating n times we can precisely locate any 3 x n bit number in 3 dimensional space.

The rest of this paper is organized as follows: In section 2 we describe our visualization techniques and what information each of them can provide; Section 3 describes some features of the VAST interface; In section 4 we provide illustrative examples of how our techniques can be applied to extract useful information from raw BGP datasets; Section 5 describes some recent related work in the area of visual techniques for BGP topology analysis and how VAST complements the current data analysis and visualization toolsets; and finally section 6 provides a brief conclusion and outlines some future work.

2. THE VAST VISUALIZATIONS

VAST is a 3-dimensional graphical tool that implements a set of visualizations that can highlight and extract relevant information from raw BGP update messages. Some of these visualizations are based on a quad-tree algorithm, while others are based on a more complex octo-tree algorithm.

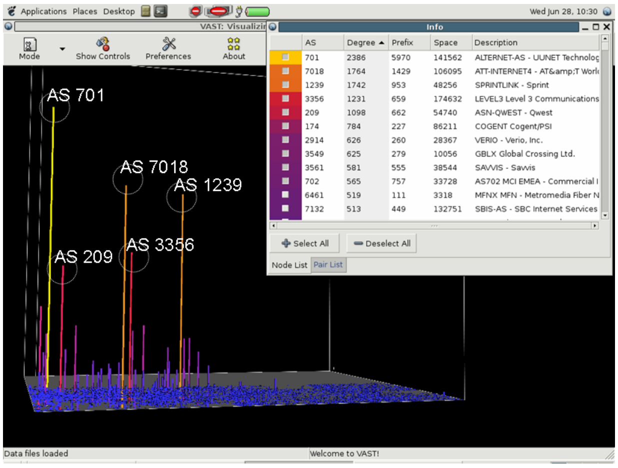
The quad-tree based visualizations represent Autonomous System Numbers (ASN) on a 2-dimensional plane. We start by taking the binary representation of the ASN and use 2 bits at a time to determine which quadrant those bits represent. Starting with a fixed size square we divide the square into four equal quadrants by cutting the x and the y axis in half. The first bit then selects top/bottom, while the second bit selects left/right. In this this way 2 bits specify a single quadrant of the original square. By repeating this algorithm n times it is possible to precisely represent any 2 x n bit number. The third dimension can then be used to express quantitative properties of each of those entities. Colors, line-widths, and size give us some addition degrees of freedom which can be used to represent other aspects of the dataset. Further details regarding the use of this technique for a visual representation of network traffic data can be found in [2].

While the quad-tree approach can be useful for representing per AS information, it is difficult to represent relationships *between* multiple Autonomous Systems (AS) with this method. The 2-dimensional nature of the algorithm limits our ability to effectively depict topological information. Therefore, while this approach can be useful to represent the distribution of the number of prefixes originated by each ASN in the Internet, it cannot easily be used to represent information about the links between each other.

In order to depict the topological relationships between AS numbers we extend the quad-tree algorithm to a third dimension. The resulting octo-tree algorithm is a method that allows us to map an ASN into 3-dimensional space instead of a 2-dimensional plane. The octo-tree algorithm uses 3 bits at a time to specify a specific sub-cube of a starting fixed size cube. Figure 1 illustrates how we can use 3 bits to specify one of eight possible sub-cubes that are generated by slicing the original cube by half in terms of length, width, as well as height. Repeating this algorithm n times we can uniquely locate any 3 x n bit number within a fixed size cube. Once again we start by using the binary representation of an AS number and use 3 bits at a time in order to determine the location of its visual representation. A minor enhancement that we use for this technique to work correctly is that we pad the 16-bit ASN by adding two zeroed bits at the end to get a total of 18 bits. Repeating the octo-tree algorithm 6 times gives us the precise location of this number in a 3-dimensional space. Relationships between various pairs of ASN can now be easily depicted. Below we briefly describe three quad-tree and four octo-tree based visualizations. that are implemented in the VAST tool.

2.1 Quad-Tree Visualizations

The quad-tree based visualizations are most effective for expressing information about *individual* AS numbers. The quad-tree algorithm maps an AS number onto a 2-dimensional square at the base of the cube, the third dimension of height is then used to represent the relative values of various metrics. The techniques we used here are similar in nature to the core platform used in [2], where they were used to represent IP addresses from netflow records.



	AS	Degree ▲	Prefix	Space	Description
□	701	2386	5970	141562	ALTERNET-AS - UUNET Technolog
□	7018	1764	1429	106095	ATT-INTERNET4 - AT&amp;T Worl
□	1239	1742	953	48256	SPRINTLINK - Sprint
□	3356	1231	659	174632	LEVEL3 Level 3 Communications
□	209	1098	662	54740	ASN-QWEST - Qwest
□	174	784	227	86211	COGENT Cogent/PSI
□	2914	626	260	28367	VERIO - Verio, Inc.
□	3549	625	279	10056	GBLX Global Crossing Ltd.
□	3561	581	555	38544	SAVVIS - Savvis
□	702	565	757	33728	AS702 MCI EMEA - Commercial I
□	6461	519	111	3318	MFNX MFN - Metromedia Fiber N
□	7132	513	449	132751	SBIS-AS - SBC Internet Services

Figure 2: Out-degree for each observed Autonomous System Number: This figure clearly shows how very few AS numbers have a large number of peering sessions

2.1.1 *ASN Out-degree*

This visualization is a simple visual representation of the relative out-degree values of different AS numbers. For each AS the height of its bar is based on its out-degree. In addition, a gradient of colors is used to emphasize the scale of the out-degree. The color gradient ranges from blue for AS numbers that have low values of out-degree, to red that is used to depict AS numbers that have medium out-degree values, to yellow which is used for high out-degree ASes. Large out-degree ASes are important as they are generally considered to form the core of the Internet infrastructure. Using this visualization, it is easy to quickly identify the AS numbers with the highest out-degree. By using this method on a dataset consisting of a single day of BGP update messages collected at the Routeviews collectors [1] we can easily generate a view that illustrates the out-degree value for each AS in that dataset. An example of this visualization can be see in Figure 2. This helps provide a visual understanding of the well known fact that only very few core ASes in the Internet have high levels of connectivity.

2.1.2 *Unique Prefix Originations per AS*

This visualization is used to depict the number of unique prefixes that are being originated by any particular AS number. For each AS the height of the bar represents the number of unique prefixes it was seen to originate in the input dataset. The larger the number of unique prefixes that are originated by an AS, the greater the height of that AS on the z-axis. Using this method we can easily identify AS numbers that are advertising large numbers of prefixes. From a security perspective this information can be important as an outage in one of the larger ASes would impact a large portion of the Internet. Similarly, unusual or strange AS numbers that are suddenly seen to advertise large number of prefixes can be an easily spotted as potential sources of wide-spread BGP instability caused by configuration errors or perhaps even malicious attack. We provide an example of this in Section 4.1.

2.1.3 *Address Space Announced per AS*

In addition to the number of prefix originations, the size of the address space being advertised by different AS numbers is also important. In this visualization we use the z-axis to represent the total amount of address space that is being announced by different AS numbers. Once again this information is of importance from a security and planning perspective as it illustrates what the impact of an outage in one of the larger core Autonomous Systems might be. It is also interesting to note that while this visualization also exhibits the same features as the out-degree visualization regarding only a small fraction of the AS numbers showing a large values, the AS numbers that exhibit these large values are not generally the same. Once again, this provides a visual basis for our intuitive understanding that the core Internet Autonomous Systems are focused on maintaining a large peering set, and often delegate the origination of addresses to their downstream customers. Section 4.3 provides an example which illustrates the use of this visualization.

2.2 Octo-Tree Visualizations

As described earlier the octo-tree algorithm allows us to depict AS numbers as points in 3-dimensional space. This technique can then be used to visualize various AS topological relationships. In the following subsections we describe 5 different views that can be used to extract or highlight different properties embedded in the input dataset. In our visualizations the use of color and node size remain consistent, while the interpretation of line-width changes with the view. The size of a node is scaled relative to its out-degree or the number of peering sessions it is seen to have in the dataset. As the out-degree of an AS increases its color ranges from blue to red to bright yellow. This ensures that these AS numbers are easily spotted. ASes with a high out-degree will appear as large yellow cubes, while ASes with small out-degree will be smaller blue cubes. Lines are then drawn between connected AS numbers to show topological links. The thickness of these lines is scaled based on a different metric for each of the views described below.

2.2.1 AS Connectivity - No Scaling

The first visualization is a simple representation of all the ASes that have been seen in the input BGP dataset. Lines are drawn between AS nodes to show the presence of a topological link but the line width is kept uniform for all links. This visualization is well suited to simply explore the connectivity of a particular AS. Selecting an AS number from the VAST information window will display all observed peering sessions of that node.

2.2.2 Out-degree

The most commonly explored aspect of AS topology data is the degree of various ASes. The out-degree of an AS is defined as the number of ASes that a given AS is directly connected with. VAST can easily explore and highlight the out-degree of various ASes by scaling the line thickness of each of the peering sessions of any AS by the out-degree of its peer. Figure 4 shows an example where VAST is used to explore the connectivity of the SWITCH network(AS 23005). The figure shows its observed connectivity based on a BGP update messages dataset from August 18th 2006. In this view, the link between SWITCH and AS 701 (UUNET), is clearly highlighted by a much thicker line. This peering link is relatively more important as AS 701 has a much higher out-degree than any of the other peering links.

2.2.3 Unique Prefixes per AS Pair

In this visualization we vary the line thickness of the lines connecting AS numbers based on the number of unique prefixes that are being advertised over a particular pair of ASes. This can be used to highlight and draw attention to AS links that are used by a large number of prefixes and are potentially more important. We illustrate the use of this view in Section 4.1.

2.2.4 AS Pair Frequency

This view attempts to highlight pairs of ASes that are seen more frequently in ASPATH values in BGP update messages than others. The AS pairs that are seen more frequently in the BGP update messages dataset is highlighted by drawing a thicker line connecting them. A high frequency of an AS pair may indicate instability of the link, or it might be an indicator of paths or links that are used by a large number of routing messages and are therefore more important.

2.2.5 Address Space per AS Pair

In addition to the number of unique prefixes advertised over a particular link, the aggregate IP address space advertised is also an important metric. In order to highlight particularly important links we vary the line thickness of the line connecting any pair of AS numbers based on the aggregated address space advertised over each link. Section 4.4 provides an example of how this view can be useful for BGP data analysis.

3. THE VAST INTERFACE

One of the key features of VAST is the ability to explore and manipulate the data being visualized in order to extract relevant information.

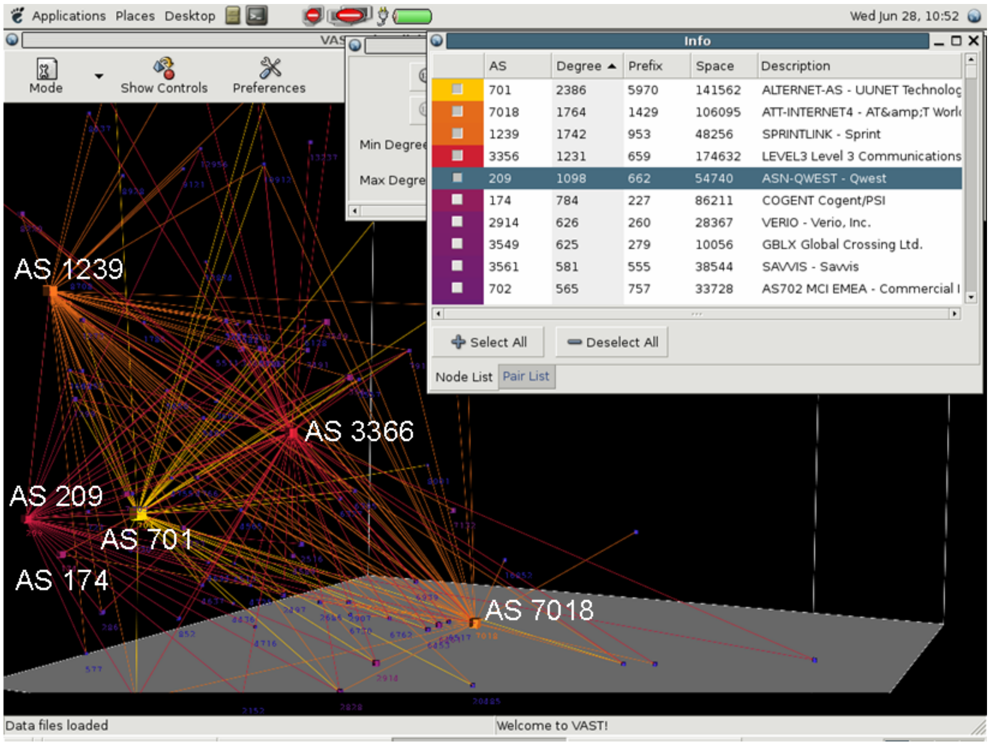
3.1 OpenGL View

The VAST interface uses OpenGL to visualize data from raw BGP messages. Data can be loaded from Routeviews or any BGP data source via libbgpdump. The interface provides the ability to navigate easily through the 3d OpenGL view via simple mouse controls. Left-click is used to rotate, middle-click to zoom in/out, and right-click to pan left/right/up/down. By navigating through the 3d view, one can focus attention to specific areas of the display that are of interest.

3.2 Information Window

The information window gives quantitative details that complement the graphical visualizations. The window lists statistics such as ASN, AS description, degree, announced unique prefixes, and announced address space. The entries of the list can be sorted ascending or descending for each statistic, exposing potentially anomalous statistics. The color column is especially useful for correlating the data in the list with the main visualization.

The information window also provides a checkbox for each AS in the list. These checkboxes affect the octo-tree visualizations and allow an operator to show/hide the details of



	AS	Degree ▲	Prefix	Space	Description
□	701	2386	5970	141562	ALTERNET-AS - UUNET Technolog
□	7018	1764	1429	106095	ATT-INTERNET4 - AT&amp;T Worl
□	1239	1742	953	48256	SPRINTLINK - Sprint
□	3356	1231	659	174632	LEVEL3 Level 3 Communications
□	209	1098	662	54740	ASN-QWEST - Qwest
□	174	784	227	86211	COGENT Cogent/PSI
□	2914	626	260	28367	VERIO - Verio, Inc.
□	3549	625	279	10056	GBLX Global Crossing Ltd.
□	3561	581	555	38544	SAVVIS - Savvis
□	702	565	757	33728	AS702 MCI EMEA - Commercial I

Figure 3: Visualizing the Internet Core: This figure depicts the AS topological connectivity of the top 5 most connected nodes in the Internet. For clarity, only links with ASes that themselves have at least 50 peering sessions are shown

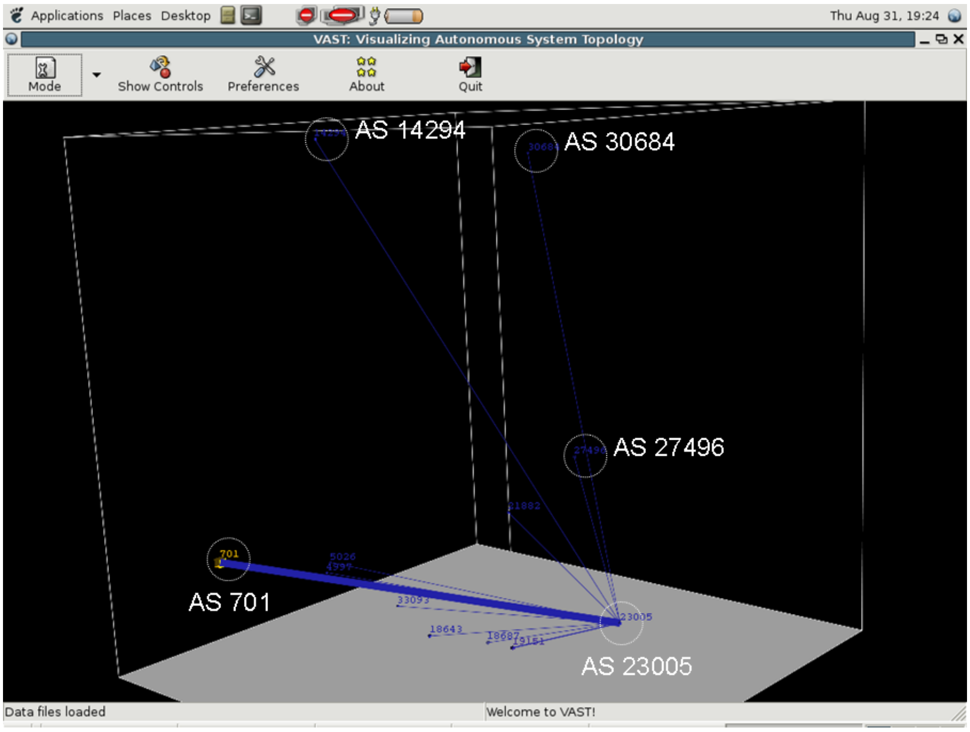
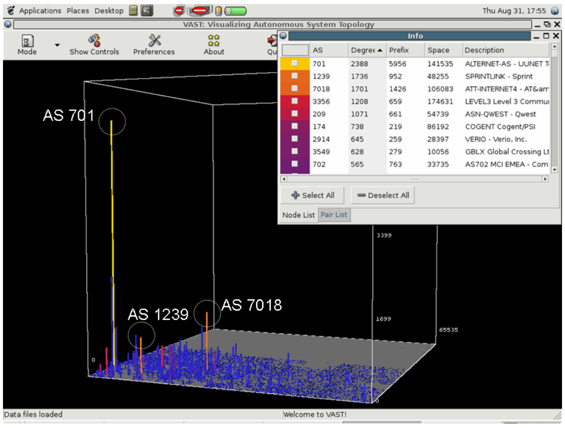


Figure 4: Visualizing AS Topology: This figure depicts the AS topological connectivity of the SWITCH Autonomous System as observed on Aug 18th 2006. The link with AS 701 is clearly highlighted in the visualization drawing attention to the key peering link that connects SWITCH to the Internet core.



	AS	Degree	Prefix	Space	Description
□	701	2388	5956	141535	ALTERNET-AS - UUNET T
□	1239	1736	952	48255	SPRINTLINK - Sprint
□	7018	1701	1426	106083	ATT-INTERNET4 - AT&am
□	3356	1208	659	174631	LEVEL3 Level 3 Commu
□	209	1071	661	54739	ASN-QWEST - Qwest
□	174	738	219	86192	COGENT Cogent/PSI
□	2914	645	259	28397	VERIO - Verio, Inc.
□	3549	628	279	10056	GBLX Global Crossing Lt
□	702	565	763	33735	AS702 MCI EMEA - Com

Figure 5: AS 9121 Route Leakage Incident: This figure shows the number of prefixes originated by various AS numbers on December 22nd 2004

specific ASes in the OpenGL view. For example, enabling the checkbox for AS 701 will render lines from the AS 701 node to all the ASes it's peered with and the connecting lines will be scaled based on which visualization mode is currently active. "Select All" and "Deselect All" buttons are present to make it simple to view the relationships between all the AS nodes.

3.3 Controls Window

The primary function of the controls window is to allow an operator to tune thresholds based on several metrics to control what data is being displayed in the 3d view. These slider-bar threshold controls can easily filter out unnecessary data and allow slicing and dicing of the data to more efficiently extract pertinent information from the visualization.

For example, if the "Select All" button is enabled as described in the previous section, a very large number of lines connecting all the AS nodes will be rendered and the threshold controls play an invaluable role in managing the large number of interconnections.

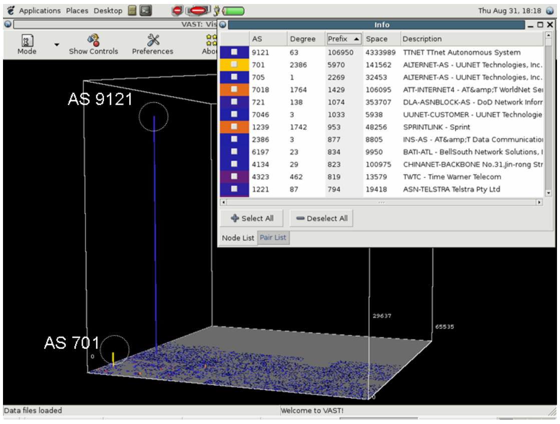
A checkbox is also present to enable/disable text labels that are rendered in the 3d view to provide easy identification of the AS nodes. As enabling the labels adds a lot of information to the display, it is meant to be used in conjunction with the threshold controls.

4. ILLUSTRATIVE EXAMPLES AND DISCUSSION

4.1 Route Leakage Event - 12/24/2004

A well known routing anomaly that is commonly used as a case study in BGP routing research is the AS 9121 route leakage incident. On December 24th 2004, AS 9121 started re-announcing a large number of prefixes into the Internet. This resulted in some portions of the Internet being unreachable for the duration of the anomaly. The problem was quickly identified and resolved.

Using VAST we can run a quick analysis on the raw data archives for this time period and obtain a visual understanding of this event. For the analysis below, we use two datasets. The first is for December 22nd 2004, this serves as our baseline. The second is for the day of the incident, December 24th 2004. Figures 7 and 8 present a side-by-side comparison of the connectivity of AS 9121 and AS 6762 on those days. The line thickness in these visualizations represents the number of prefixes being announced over a particular link. Figure 8 clearly shows that on the day of the incident, AS 9121 is seen to announce a very large number of prefixes over its link with AS 6762. The line thickness of the link between AS 9121 and 6762 changes dramatically between the December 22nd and the December 24th datasets. Moreover, AS 6762 re-announces these into the core of the Internet via its link with AS 701. Using VAST we can quickly obtain information not only about the origin of the anomaly, but also where security mechanisms were inadequate resulting in its spread. Not only do we see AS 9121 create the anomaly by announcing large numbers of prefixes, but we also see that AS 6762 failed to properly detect and prevent their propagation via the use of filtering mechanisms.



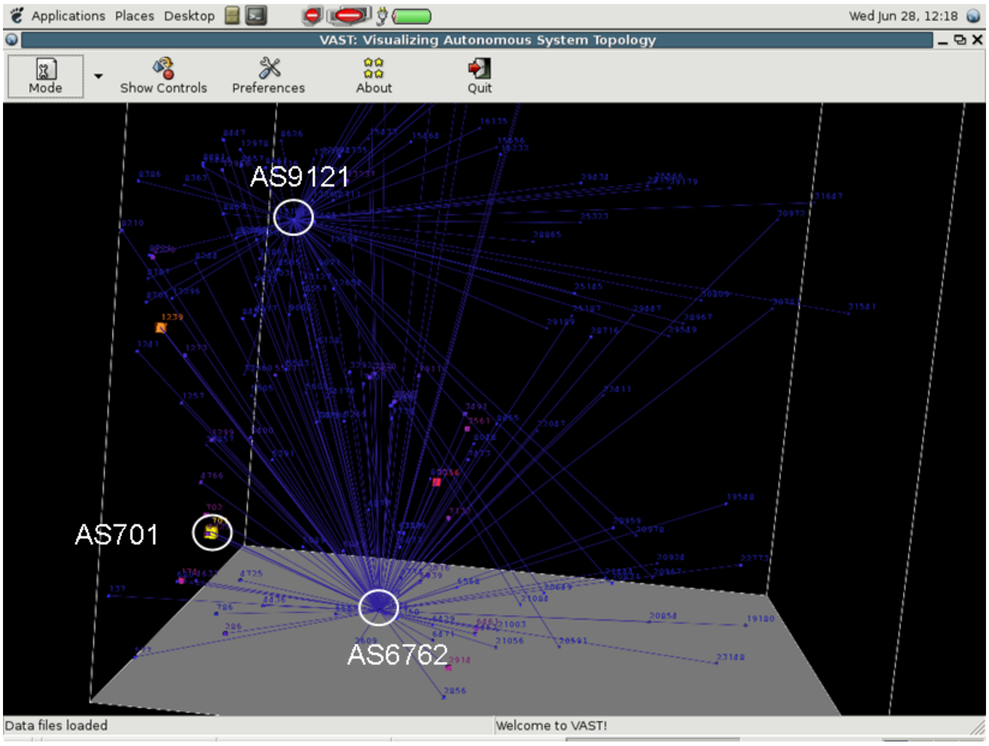
	AS	Degree	Prefix	Space	Description
□	9121	63	106950	4333989	TTNET TTnet Autonomous System
□	701	2386	5970	141562	ALTERNET-AS - UUNET Technologies, Inc.
□	705	1	2269	32453	ALTERNET-AS - UUNET Technologies, Inc.
□	7018	1764	1429	106095	ATT-INTERNET4 - AT&amp;T WorldNet Se
□	721	138	1074	353707	DLA-ASNBLOCK-AS - DoD Network Infor
□	7046	3	1033	5938	UUNET-CUSTOMER - UUNET Technologie
□	1239	1742	953	48256	SPRINTLINK - Sprint
□	2386	3	877	8805	INS-AS - AT&amp;T Data Communicatio
□	6197	23	834	9950	BATI-ATL - BellSouth Network Solutions, I
□	4134	29	823	100975	CHINANET-BACKBONE No.31,Jin-rong Str
□	4323	462	819	13579	TWTC - Time Warner Telecom
□	1221	87	794	19418	ASN-TELSTRA Telstra Pty Ltd

Figure 6: AS 9121 Route Leakage Incident: This figure shows the number of prefixes originated by various AS numbers on December 24th 2004. The number of unique originations by AS9121 dwarfs all other AS numbers.

Figure 5 shows another view of the same dataset. This figure shows the relative number of unique prefixes announced by various AS numbers on December 22nd 2004. Figure 6 show the same view for the December 24 2004 dataset. Once again the AS 9121 anomalous behavior is clearly visible as the large spike in the number of unique prefixes originated by AS 9121.

4.2 Identifying Critical Internet Infrastructure

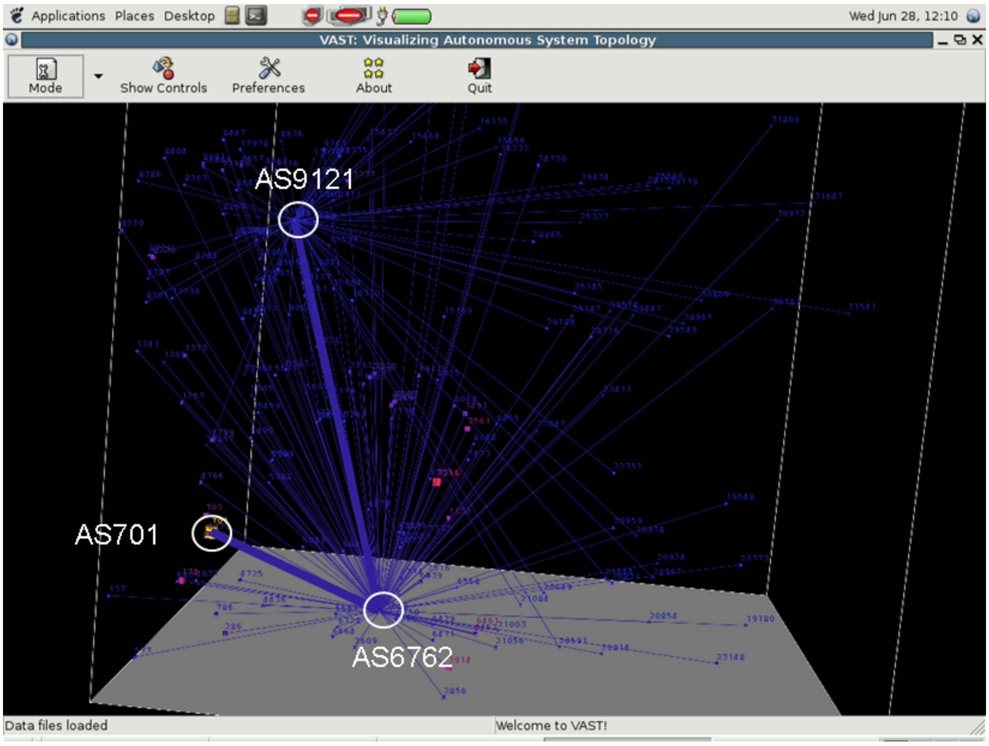
One relatively straight forward use of VAST is in quickly identifying critical Internet infrastructure based on historical BGP data. After we have loaded in a BGP raw dataset into VAST, we can easily use the slider bar controls in concert with the information window to identify portions of the Internet AS topology that are most important for its operation. For example, Figure 3 shows the AS level connectivity of the top five most connected nodes in the Internet. For


Applications Places Desktop
Wed Jun 28, 12:18
VAST: Visualizing Autonomous System Topology
Mode
Show Controls
Preferences
About
Quit
AS9121
AS701
AS6762

Data files loaded

Welcome to VAST!

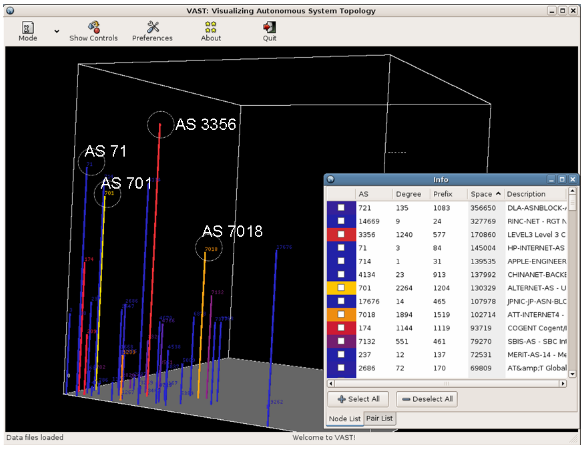
Figure 7: AS 9121 Route Leakage Incident: This figure shows the AS level connectivity between AS 9121 and AS 6762 on December 22nd 2004


Applications Places Desktop
Wed Jun 28, 12:10
VAST: Visualizing Autonomous System Topology
Mode
Show Controls
Preferences
About
Quit
AS9121
AS701
AS6762

Data files loaded

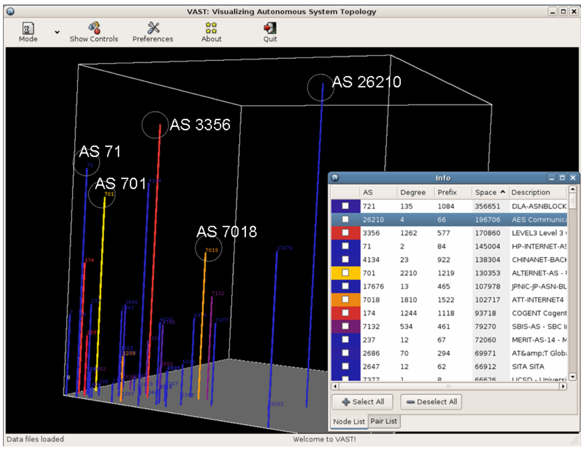
Welcome to VAST!

Figure 8: AS 9121 Route Leakage Incident: On December 24th 2004 there is a dramatic change in the connectivity between AS 9121 and AS 6762. A very large number of prefixes are now being announced over that link. AS 6762 is in turn passing them onto AS 701.



	AS	Degree	Prefix	Space ^	Description
□	721	135	1083	356650	DLA-ASNBLOCK-/
□	14669	9	24	327769	RINC-NET - RGT N
□	3356	1240	577	170860	LEVEL3 Level 3 C
□	71	3	84	145004	HP-INTERNET-AS
□	714	1	31	139535	APPLE-ENGINEER
□	4134	23	913	137992	CHINANET-BACKE
□	701	2264	1204	130329	ALTERNET-AS - U
□	17676	14	465	107978	JPNIC-JP-ASN-BLC
□	7018	1894	1519	102714	ATT-INTERNET4 -
□	174	1144	1119	93719	COGENT Cogent/I
□	7132	551	461	79270	SBIS-AS - SBC Int
□	237	12	137	72531	MERIT-AS-14 - Me
□	2686	72	170	69809	AT&amp;T Global

Figure 9: Autonomous Systems announcing large amounts of address space on September 8th, 2005



	AS	Degree	Prefix	Space ^	Description
□	721	135	1084	356651	DLA-ASNBLOCK
□	26210	4	66	196706	AES Communica
□	3356	1262	577	170860	LEVEL3 Level 3 (
□	71	2	84	145004	HP-INTERNET-AS
□	4134	23	922	138304	CHINANET-BACK
□	701	2210	1219	130353	ALTERNET-AS - I
□	17676	13	465	107978	JPNIC-JP-ASN-BL
□	7018	1810	1522	102717	ATT-INTERNET4
□	174	1244	1118	93718	COGENT Cogent
□	7132	534	461	79270	SBIS-AS - SBC I
□	237	12	67	72060	MERIT-AS-14 - M
□	2686	70	294	69971	AT&amp;T Glob:
□	2647	12	62	66912	SITA SITA
□	7377	1	8	66626	UCSD - Universi

Figure 10: Autonomous Systems announcing large amounts of address space on September 9th, 2005. While the values of the other Autonomous Systems remain roughly the same, AS 26210 is suddenly seen to originate a large amount of address space indicating a possible anomaly.

clarity, only links where the AS has more than 50 connections are shown. Therefore this figure is a representation of the most important nodes in the Internet topology from a connectivity perspective. Similarly we can identify links or AS pairs that are relatively more important from the perspective of number of unique prefixes carried over them or even the total amount of address space that is advertised over them.

By using VAST it is relatively simple to obtain a list of the top 10 most connected AS numbers in the Internet, by simply loading in the raw BGP update messages dataset and selecting the "Out-degree per AS" view. By selecting to view the "Out-degree Topology" view and using the check boxes in the information window to select the top 10 most connected AS numbers, we can visualize the connectivity of the Internet core in a 3-dimensional space. The zoom/rotate/pan mouse controls further enhance our ability to truly explore this topology.

4.3 Visualizing an Address Space Hijacking Incident

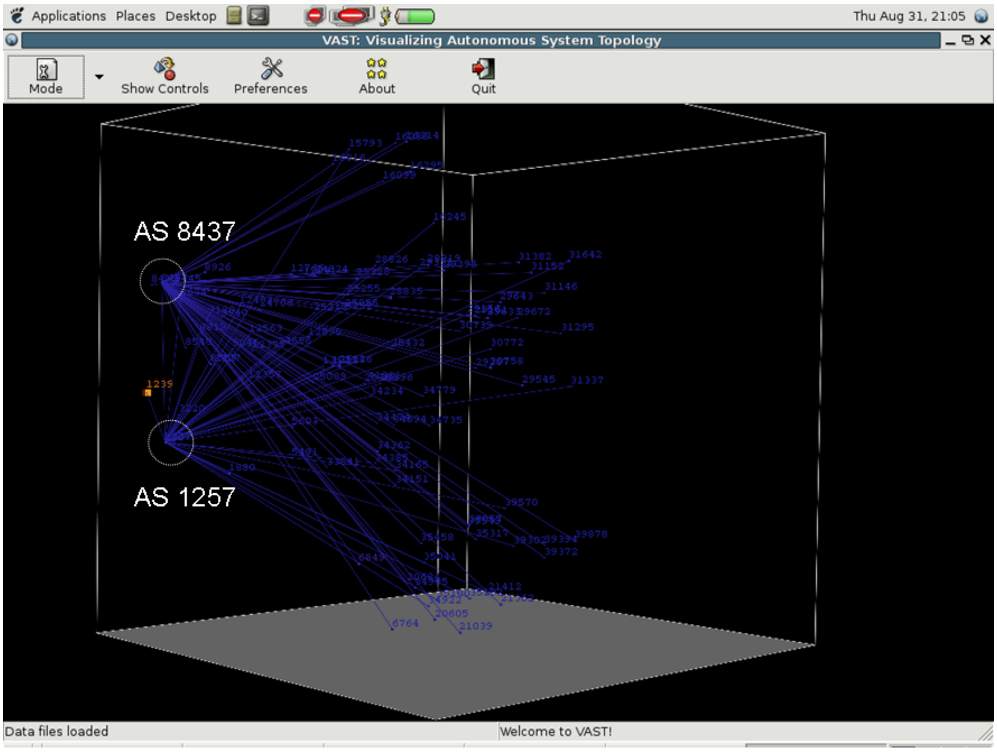
In this example we focus our attention on the AT&T 12/8 prefix hijack event that occurred on September 9th, 2005 [3]. During this routing anomaly AS 26210 inadvertently announced the 12/8 prefix which belongs to AT&T. We examined data from two datasets for this event. The first dataset serves as our baseline and is for September 8th, 2005. The second dataset is the raw data from Routeviews for September 9th 2005. Figures 9 and 10 show the "Address-space-per-AS" view for each of these datasets. Figure 9 shows the various AS numbers that are advertising the most address space on our baseline day. In Figure 10 we see that though the address space being advertised by the other AS numbers stays relatively the same, AS 26210 suddenly appears to advertise a large address space. The information window shows the same information, and this quickly tells that AS 26210 is a small service provider in Bolivia, and therefore a sudden large announcement of address space is most likely an anomaly. Once an anomaly has been quickly identified visually, we can obtain further detailed information about it by running specific queries on indexed datasets [4].

4.4 Bogon Route Leak - 08/18/2006

On August 18th 2006 it was noticed that AS 8437 was announcing a large percentage of the unused IP address space [5]. In order to analyze data for this event we use the "Address Space per AS pair" view described in Section 2.2.5. Figure 11 shows the connectivity between AS 8437 and AS 1257 on August 17th 2006. Figure 12 shows the connectivity of these two ASes on August 18th 2006. Compared to Figure 11 this figure clearly shows a thick anomalous line connecting these two ASes indicating that a much larger address space is being advertised between these ASes on August 18th 2006. These figures also indicate that had AS 1257 did not have the proper filters in place to mitigate this misconfiguration. Using the alternate view into the data provided by the "Address Space per AS" method described in Section 2.1.3 confirms that AS 8437 is indeed originating an unusually large amount of address space. Once we have visually identified an anomaly such as this we can obtain detailed information about the prefix announcements by this AS using tools such as BGP-Inspect [4].

5. RELATED WORK

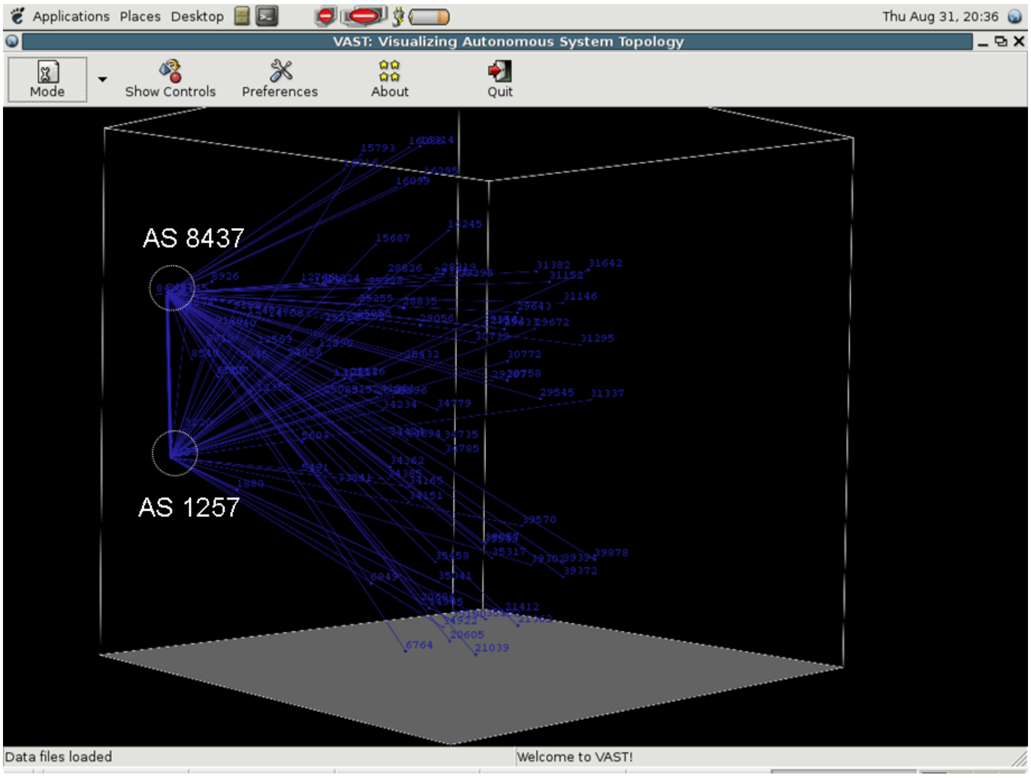
The most well known example of a visual representation of the AS topology of the Internet is the AS Level Internet Graph [6] designed by the Cooperative Association for Internet Data Analysis (CAIDA). This visualization depicts the AS topology in polar coordinates by using out-degree of an AS to determine the distance from the center of a circle (radius) and its geographic location to determine its position around the circle (theta). This technique clearly illustrates some simple features of the Internet AS topology, and provides insights into peering richness coupled with geographic information. However, while this visualization technique provides a static image, VAST provides an interactive tool that can be used to explore and manipulate the topological data. VAST is based on 3-dimensional variations of quad-tree/octo-tree algorithms, leading to a more intuitive visual representations. In addition, VAST provides a


Applications Places Desktop
Thu Aug 31, 21:05
VAST: Visualizing Autonomous System Topology
Mode
Show Controls
Preferences
About
Quit
AS 8437
AS 1257

Data files loaded

Welcome to VAST!

Figure 11: Autonomous Systems announcing large amounts of address space on August 17th, 2006. A thin line connects AS 8437 and AS 1257 indicating a normal amount of address space is being advertised over that link.


Applications Places Desktop
Thu Aug 31, 20:36
VAST: Visualizing Autonomous System Topology
Mode
Show Controls
Preferences
About
Quit
AS 8437
AS 1257

Data files loaded

Welcome to VAST!

Figure 12: Autonomous Systems announcing large amounts of address space on August 18th, 2006. The thick line connecting AS 8437 and AS 1257 indicates that a large amount of IP address space was advertised over that link.

sequence of visualizations that each attempt to highlight a different aspect of the raw BGP data.

In addition to the AS Level Internet Graph work from CAIDA there are a number of other analysis tools that attempt to provide insight into raw BGP data. LinkRank [7] is a graphical tool that allows users to playback archived BGP data. This can be used to detect routing changes. The number of routes over a particular pair of AS numbers is the key metric that is used in order to determine the rank of a link. BGPlay [8] is also a BGP data playback that allows users to visualize BGP routing dynamics. Java Autonomous System Path Visualization Visualization Interface (Jaspvi) [9] is a similar tool that allows users to visualize the AS path connectivity in a 2-dimensional space. ASExplorer [10] was also a tool that could represent BGP routing dynamics and AS link relationships visually. The quad-tree visualization method has also been used previously for BGP anomaly detection[11]. [2] describes various techniques that can be used to visualize live netflow traffic. Each of the tools described above has its own unique features that focuses on a single particular aspect of the BGP data. By comparison VAST provides a much more comprehensive set of visualizations that can highlight different aspects of a dataset.

VAST differs from most of the related tools described above as it does not provide BGP data playback capability. Instead VAST presents a series of visualizations of data summaries extracted from the raw BGP data in an interactive 3-dimensional space. Moreover, most of the other tools are only able to visually represent limited portions of the entire AS number space and have limited interactive capabilities. VAST has the ability to represent the entire 16-bit AS number space in a single visualization providing users with a single comprehensive view which can be easily manipulated.

6. CONCLUSIONS AND FUTURE WORK

In this paper we have presented a set visualization techniques that can be used to easily extract relevant information from large amounts of raw BGP data. We have implemented these techniques to build a unique AS topology visualization tool called VAST. The basis for VAST are modified versions of the quad-tree and octo-tree algorithms that allow us to represent Autonomous System Numbers in a fixed amount of 3-dimensional visualization space. In addition, we make use of node size scaling, colors, and line width scaling to highlight different aspects of Internet AS topology. Slider bars and other controls allow an operator to filter out information that may not be relevant. VAST implements a series of visualizations based on the raw input BGP data, each of which attempts to highlight a different metric from the dataset. With the help of some well known routing anomaly incidents we have shown how VAST is a useful addition to the current set of BGP data analysis tools. VAST can be a useful tool for performing BGP routing forensics, visual anomaly detection and representation as well as the identification of critical routing infrastructure.

While VAST currently implements a set of different views, we are investigating how we can add additional visualizations that can be used to represent other important features of the raw BGP data. Currently, VAST uses information from all the BGP peers in the input BGP datasets to construct BGP topology. One interesting feature that we would like to add is the ability to specify that the AS topology information only be based on specific peering sessions. This would essentially allow us to compare different "views" of the Internet AS topology. Additionally, we are examining the feasibility of adapting VAST in order to analyze data from a live BGP peering session.

Acknowledgments

The work described in this paper was supported in part by the Department of Homeland Security (DHS) under contract number NBCHC040146.

7. REFERENCES

[1] University of Oregon. Route views project. *http://www.routeviews.org/*, January 2005.

[2] J. Oberheide, M. Goff, and M. Karir. Flamingo: Visualizing internet traffic. *Proceedings of IEEE/IFIP Network Operations & Management Symposium (NOMS)*, April 2006.

[3] Merit Network Inc. BGP-Inspect-Routeviews. *http://bgpinspect.merit.edu/reports.php*, July 2006.

[4] D. Blazakis, M. Karir, and J.S. Baras. BGP-Inspect: Extracting Information from Raw BGP Data. *Proceedings of IEEE/IFIP Network Operations & Management Symposium (NOMS)*, April 2006.

[5] J. Karlin. NANOG Mailing List. *http://merit.edu/mail.archives/nanog/msg01700.html*, Aug 2006.

[6] Cooperative Association for Internet Data Analysis (CAIDA). Visualizing internet topology at a macroscopic scale. *http://caida.org/analysis/topology/as_core_network*, April 2005.

[7] M. Lad, D. Massey, and L. Zhang. Link-rank: A graphical tool for capturing bgp routing dynamics. *Proceedings of IEEE/IFIP Network Operations & Management Symposium (NOMS)*, April 2004.

[8] G. Di Battista, F. Mariani, M. Patrignani, and M. Pizzonia. Bgplay: A system for visualizing the interdomain routing evolution. *Proceedings of Graph Drawing (GD)*, 2003.

[9] M. Prodanovic and N. Micic. Java Autonomous System Path Visualization Visualization Interface (Jaspvi). *http://lab.verat.net/Jaspvi/*, May 2002.

[10] R. Malan, F. Jahanian, and S. Subramanian. Salamander: A push-based distribution substrate for internet applications. *Proceedings of the USENIX Symposium on Internet Technologies and Systems (USITS)*, December 1997.

[11] S.T. Teoh, K-L. Ma, S.F. Wu, and X. Zhao. A visual technique for internet anomaly detection. *Proceedings of the 14th IFIP/IEEE Workshop on Distributed Systems: Operations and Management*, October 2003.