Visualizing Threats: Improved Cyber Security Through Network Visualization

Intended audience

This white paper has been written for anyone interested in enhancing an organizational cyber security regime using advanced network visualization.

It is not intended as a full or technical introduction to either subject, but instead provides examples of how network visualization has been deployed by KeyLines customers as part of a wider cyber security effort.

For more information

To learn more about any of the examples contained in this white paper, or to get a free demonstration of the KeyLines network visualization toolkit, just get in touch:

http://keylines.com/contact.

Contents

Introduction ........................................................................................................................................2
Making better use of your data ..........................................................................................................2
Case 1: Understanding unusual network traffic ..............................................................................3
Case 2: Deconstructing a botnet network .........................................................................................23
Case 3: Unusual login habits ..........................................................................................................24
Case 4: Analysing malware propagation .........................................................................................27
Case 5: Adding a Geographic and temporal element .....................................................................29
What is KeyLines? .........................................................................................................................31
Introduction

Cyber security is a central concern for all organizations. The iterative (and often ad-hoc) approach to building and maintaining IT networks has left them with in-built and growing vulnerabilities.

Criminals, terrorists, nation states, as well as activists and opportunistic amateurs, pose a real and persistent threat to corporate and government IT systems – a situation exacerbated by their complexity. Cyber security professionals face an almost impossible task on three fronts:

- Proactively maintain a secure network perimeter
- Constantly monitor for new and emerging threats
- Perform forensics on previous attacks.

Making better use of your data

Cyber security is a data-led battle, with terabytes of disparate information collated from around the enterprise into centralized dashboards, sometimes in dedicated Network Operations Centers (NOCs) or Security Operations Centers (SOCs).

Existing tools do an excellent job of collating this data, often with automated real-time alerts. Too frequently, however, they lack the visualization capability required for humans to interpret the data. This means alerts are not effectively investigated and post-attack forensics are inefficiently managed.

Integrating network visualization into the enterprise cyber security dashboard is a logical and effective approach to the manual investigation of real-time alerts and post-attack forensics.

The benefits of deploying network visualization as part of a complete cyber security management platform include:

1. **Maintenance of a secure perimeter**
   
   Quickly and efficiently interpret large volumes of real-time network data, leading to faster and more informed decisions when threats are detected.

2. **Effectively understand emerging threats**
   
   Cyber attackers’ tactics change constantly. The added context provided by network visualization helps uncover vulnerabilities before they are exploited.

3. **Perform efficient post-attack forensics**
   
   Network visualization gives a 360-degree view of network data, providing the context required to fully understand past attacks.
Case 1: Unusual network traffic

Understanding where traffic is going outside of your network is a vital task. Once an attacker has compromised your network often the next step is to ‘call home’ in order to accept the tasks requested by the hacker, e.g. sending spam or collecting personal details, or simply to export data stolen from your system.

Similarly, high levels of inbound traffic should raise concerns, often indicating a large number of connection attempts and an imminent compromise.

This screenshot shows a filtered view of IP connections on one corporate network. The traffic volume is indicated by the weight of the link (the width of the arrow) and the node color denotes its location – green is part of the network, red is external.

Nodes are clustered using a structural layout, so nodes with similar characteristics are grouped together.

We can clearly see one external node sending a high volume of traffic – potentially requests – into the corporate network. Certain nodes are also sending high volumes externally.

There could be rational explanations for both scenarios, or they could indicate an automated attack or internal compromise.

By visualizing this data a cyber security professional can observe normal and abnormal patterns, giving a strong starting point for further investigation.
Case 2: Deconstructing a botnet network

Botnets are a common cause of unusual traffic patterns, and indicate machines in your network may be participating in DDoS attacks or spam campaigns.

Understanding the scale of your problem and finding all of the infected machines in a huge network can be challenging. Equally, understanding whether they are related or working together is virtually impossible without effective data visualization.

This first screenshot shows an unfiltered cut of traffic data taken from a university IT network: http://bit.ly/1gk7Kl1

Nodes are IP addresses; the links are packets of data. Nodes are sized by the number of packets they are processing (inbound or out) and links are weighted by packet size. Links are also colored by protocol type.

The unfiltered view, like the one shown here, is difficult to interpret.
This screenshot shows the same data but filtered to show just the sub-networks around the two most active machines. This quickly helps us cut out the noise that we do not currently need to see. This is a quick and intuitive way to find machines in a network that have potentially been taken over by a botnet system.
Case 3: Unusual login habits

Another significant indicator of compromise is network logins coming from anomalous locations.

For example, a UK-based organization with staff located only in Europe would not expect any significant volume of logins from outside that region. Likewise, most users should only access their account from a relatively narrow range of IP addresses.

With techniques more commonly used to detect fraud in transactional data, it’s possible to find suspicious login behavior by visualizing data as a network. This screenshot shows an overview of logins to an online currency exchange system.

The central node of each structure indicates an online account; each connected node is an IP address used to access that account.

We can see that most accounts have been accessed by 1-4 different IP addresses. Specific ‘star’ structures stand out. These should be investigated as they indicate an account has been accessed from many different locations.
Here we have zoomed in on two ‘star’ structures. At this level, we can see more detail:

- Green nodes indicate the user
- Purple nodes indicate the account
- Yellow nodes indicate the machines from which that ‘user’ has accessed their account.

Looking closer still, we can see more detail. The account node uses a glyph to indicate the country of registration for the account.

The node connected by a thick yellow link is the account’s ‘original’ IP address.
**Case 4: Malware propagation**

The fourth and, perhaps, most versatile cyber security use case for network visualization is to understand the propagation of malware. This model can be used for multiple activities, including:

- Understanding how far a known virus has spread and identify compromised machines
- Modeling the threat posed by malware
- Tracking offensive hacking activity
- Monitoring honeypot traps

This series of screenshots show the propagation of malware through a deliberately infected computer network. Twelve machines in the network were infected to see how the traffic spread to other machines. Over 7800 machines were included in the dataset.

All network traffic in a single chart. Yellow links indicate benign traffic; red links indicate traffic with at least 1 infected packet. Already certain machines stand out as highly active.

This screenshot shows all the traffic coming from a single machine. Nodes are sized by volume of traffic, clearly showing 7 machines were disproportionately affected.
This final screenshot shows just the infected traffic in the network. We can identify the original two machines to be infected, with the ten other machines form a cluster in a tightly connected group in the top-left.

Armed with this visualization a cyber security professional can identify machines that would propagate malware quickly, or better understand the structure of a foreign infected network.
Case 5: Adding temporal and geospatial data

Finally, it is worth looking at two important dimensions to your connected data that are difficult to convey with standard node-link visualization: time and location.

Understanding the physical location of an entity and the time of an event can unlock insight into patterns and trends, helping to uncover the origin of an attack, or predict what will happen in future.

Using KeyLines Geospatial integration and the KeyLines Time Bar, it is possible to display complex time and geographic data, without losing sight of the connections.

The example to the left shows a map of Europe superimposed with the location of PCs controlled by a botnet. Using the time bar component below the map, it’s possible to filter the data by time and date and observe the evolution of the attack using the ‘play’ button.
What is KeyLines?

KeyLines, by Cambridge Intelligence, is a unique technology for visualizing, analyzing and exploring this kind of complex connected data in an intuitive and manageable way.

Using the KeyLines toolkit, developers can rapidly build high-performance network visualization web applications to deploy alongside existing cyber security dashboards. These applications turn raw connected data into powerful interactive charts, empowering users to ‘join the dots’ and discover hidden trends.

The applications built using KeyLines run in virtually any web browser and any device. A flexible architecture means they can be deployed into existing IT environments, as part of a dashboard or as standalone tools. Data can be pulled from multiple sources and interactive charts can be shared for reporting purposes.

The advanced functionality available includes:

- Easy filtering – used in examples 2 and 4
- Multiple automated layouts – used in examples 1-4
- Social network analysis (SNA) measures – degree centrality measure used in example 2
- Node and link aggregation and grouping
- Full tablet and gesture support, including iPad and smart phone compatibility
- A time bar for visualizing temporal data – used in example 5
- A mapping integration to visualize connected data with a geospatial dimension – used in example 5.

Having easy access a visualization of the organization’s cyber security status is a real benefit to any time-pressed cyber security professional.

To register for a free trial, visit http://keylines.com/try-keylines.
We are Cambridge Intelligence

We've made it our mission to help organizations understand their connected data.

To do this, we provide everything that's needed to create impressive visualization solutions that are powerful, yet simple to use.

Our award-winning software, unrivalled knowhow and expert developer services combine to make network visualization accessible and meaningful.

**Products**
KeyLines is our first product. It’s a unique software development kit for building network visualization applications.

These applications run in any web browser and on any device. They’re also highly interactive and rich in analytical functionality, making it easy for users to get the most insight out of their data.

Organizations all over the world – from startups to government agencies – have used KeyLines to solve some of the most complex data visualization challenges.

**Knowhow**
We understand that data visualization probably isn’t your organization’s core focus. It’s also something that’s very difficult to get right.

Our data visualization experts have a wealth of knowledge and experience to help you avoid the pitfalls. We can help you build and follow a proven process, adhering to data visualization and design best practices at all time.

With our help, your solution will be easy to use and valued by your audience. In short, we’ll make sure your project is a complete success.

**Services**
We offer a range of services to help your developers make good progress, fast.

Our demos, comprehensive documentation and developer review sessions give a kick-start to your project – allowing you to be productive straight away.

We can offer support for prototyping, custom application and feature development, testing and on-going maintenance, removing the barriers stopping you from building the best application and getting it to market as soon as possible.