Poster: Security for the Common Man

Sandra G. Dykes, William L. Arensman, and Ronnie L. Killough
Southwest Research Institute
{sdykes, warensman, rkillough}@swri.org

***Abstract*—Security for the Common Man (SCM) is a new approach for detecting hidden malware by identifying a legitimate reason for outbound traffic. This method is combined with an interface that encourages user involvement. Users know what they were doing online and when they were doing it. We believe this knowledge can improve automated detection systems. SCM identifies legitimate outbound sessions by establishing a causal chain of prior sessions. Malware communication stands out because there are no legitimate causal links. We use a clock-like nova visualization to show users their 24-hour online communication pattern with suspicious sessions highlighted. Additional information is provided to help users relate their view of events to the flagged sessions. We evaluated SCM by role-playing various user types (Worker, Shopper, and Teenager) to explore the causal relationships and the visualization patterns for each. The experiments uncovered a real world example that illustrates the detection of hidden malware.**

***Keywords*-malware detection, usable security, visualization.**

I. INTRODUCTION

Most host-based security tools attempt to minimize user interaction. When users are given warnings or asked for permissions, the dialogs are typically from the perspective of the tool and often poorly understood by the user. Our goal is to encourage user involvement by presenting information that is quickly comprehended and expressed in terms of human experience. Communicating at the human level is especially important in home and small office environments where users have neither the time nor expertise to manage complex security software. This work introduces a new approach for detecting hidden malware, including rootkits, and combines it with an interface designed to encourage involvement by non-experts. In keeping with its purpose, we named this approach Security for the Common Man (SCM).

SCM is a passive network monitoring approach. It detects malware by causal analysis of outbound communication sessions; that is, sessions initiated by the local host. (A session includes all bi-directional flows between two endpoints until there is an inactivity interval of ≥ *m* minutes.) Rather than looking for malware or for behavior indicative of malware, we attempt to find a cause for outbound sessions. Legitimate sessions are often triggered by forward references in prior sessions such as URL references in a web page, HTML redirections, or host lists in a peer-to-peer application. Malware sessions stand out because there is no such causal link. However the converse is not true; a session with no causal link may be legitimate. Such sessions are benign if the site is on a whitelist or if the user initiated it. The user will usually recall the session if provided information in human terms, such as time of day, activity (shopping, music, game, etc.), site name, geographic location, reputation, etc. If neither the automated system nor the user can account for a session, then the user is alerted to the possible presence of malware.

SCM displays the user's 24-hour online pattern with a clock-like nova radial visualization. The radial characteristic of the nova makes it easy to simultaneously identify overall patterns and to spot unusual events. For example, when a user logs in at the start of the day, glancing at the nova will show if there at been any unusual overnight activity.

Because our method is based on external network monitoring, it cannot be subverted by a rootkit, has no impact on performance, and is independent of the target hardware and software. Session-based analysis allows the SCM to operate in real time because it has lower storage and computational requirements than packets or flows. Moreover, sessions are a more intuitive concept for humans.

II. CAUSAL CHAINS

Causal analysis has two elements: extracting forward references from incoming packets and linking new outbound sessions to those references. When the local host initiates a session, SCM looks for the destination in a table of forward references. If the site is found or whitelisted, SCM builds a chain of dynamically linked sessions as illustrated in Fig. 1. Otherwise the session is flagged as having no known cause.

Analysis must be efficient because it occurs at line rate. For outgoing packets, we check if this is a new outbound session. If so, we build the causal chain. For incoming packets, the analysis involves the following steps:

- Packet capture
- Bi-directional flow and session assignment
- Determination of application protocol
- Content analysis to extract references
- Update of session, reference, and DNS tables.

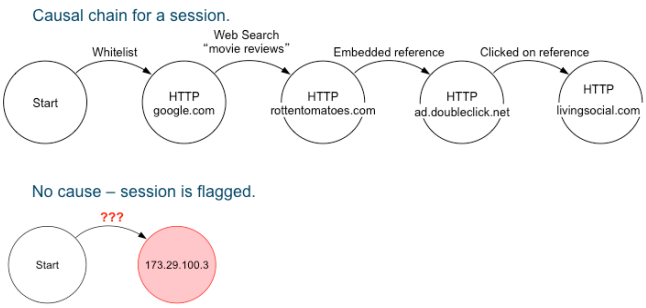

Figure 1. Causal chains for outbound sessions.

III. NOVA SESSION VISUALIZATIONS

Fig. 2 shows the nova 24-hour session visualization. Individual lines represent sessions. Line length and color denote the category (see Table 1). Angle denotes time of day, creating a 24-hour clock with midnight at the top and noon at the bottom. The clock metaphor enables users to tell at a glance when events occur. In the left graph, the empty wedge indicates when the computer was turned off. Periodic white lines are due to Chrome safe browsing updates; lack of green lines shows no online activity occurred except for a brief lunch hour period. In the right graph, the pattern indicates the computer was on all day. Green lines between 8:30 a.m. and 4:30 p.m. are consistent with workday activities, and their absence in the upper section shows that no suspicious events occurred overnight.

Highlighted sessions stand out because they are displayed with the longest lines and most vivid color. The user selects the criteria for highlighting, such as *No cause*, *Encrypted*, *Foreign country*, *Poor reputation*, or any combination thereof. Alternatively, the user may choose to highlight all sessions to a single site. SCM displays information for highlighted sessions that includes the site name, causal chain, activity (web, game, etc.), geographic location, reputation, etc. (see Fig. 3).

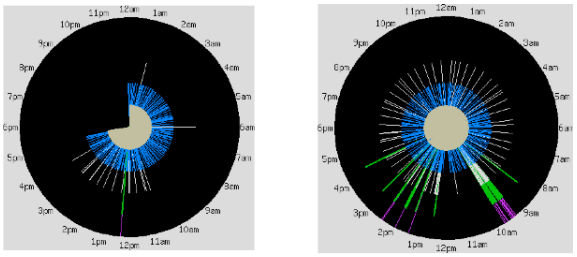

Figure 2. Nova visualization of 24-hour session activity.

TABLE 1. Session Types

Length	*Color*	*Session Type*
1	Grey	Local network management (ARP, DHCP, etc.)
2	Blue	Destination is in the local domain
3	White	Destination is on the whitelist
4	Green	Destination is a remote host not on a whitelist
5	Magenta	Highlighted session (no cause, foreign, etc.)

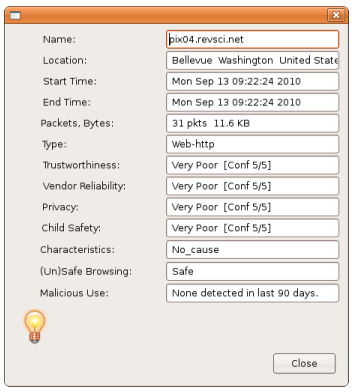

Figure 3. SCM information for a highlighted session (partial).

IV. EXPERIMENTS

Our initial testing involved role-playing three sample personas: *Worker*, *Shopper*, and *Teenager*. Worker is a professional whose computer is at the office. Shopper is a home user interested in shopping, travel, and entertainment. Teenager engages in online gaming, music, pop culture, and social networking. As shown in Fig. 4, these personas generated distinctly different patterns. Worker shows scattered online activity during working hours. Shopper shows bursts of activity during the mid-morning and late evening. Teenager shows heavy BitTorrent use overnight, even though no human was involved. SCM also reports that Teenager was online that afternoon playing World of Warcraft.

Visualizations and session details enable a user to relate flagged sessions to other knowledge. For example, mid-morning lines are suspicious if Shopper was not home or if the type of activity is unfamiliar. This is the difference between SCM and anomaly detection – whereas anomaly detection finds unusual events, SCM helps users relate an unusual event to a remembered activity.

During the experiments, SCM discovered a real-world example of hidden software. Every 15 minutes, one of our computers was connecting to c.root-servers.org, a site that provides hourly statistics for a DNS root server. As Fig. 5 shows, the nova pattern for this is striking. We identified the responsible program as MediaAgent.exe, which was installed by NBC Direct. Although NBC Direct had been uninstalled, the Media Agent software remained behind and was active for months without being detected by the user, host anti-virus tools, or enterprise security systems.

V. CONCLUSIONS

The SCM prototype demonstrates two new concepts: dynamic causal analysis of online activity and an approach for encouraging rather than discouraging user involvement in security. With further development, we believe these ideas offer a practical approach to detect malware infections. The algorithms could be developed on an embedded network device. With an inexpensive device and simple interface, this method would be feasible for users at home, in small offices, and in large enterprise organizations.

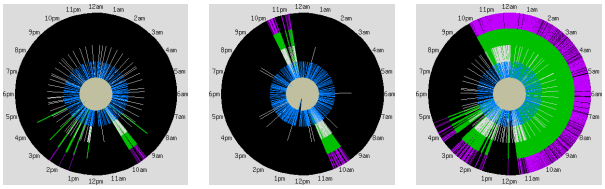

Figure 4. Results for Worker (left), Shopper (center), and Teenager (right).

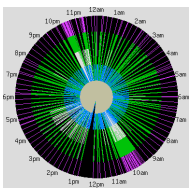

Figure 5. Discovery of hidden Media Agent software.