Human or malware? — Analysis of Web request patterns

Master thesis proposal

David Gugelmann, gugelmann@tik.ee.ethz.ch, Networked Systems Group (NSG)

Background and motivation

Web traffic has become an ideal communication channel for attackers, mainly because of two reasons. Firstly, even security-aware organizations, which block most communication protocols at their network perimeter, often permit outgoing Web traffic such that their employees can use the Web. Secondly, Web sites rely more and more on third party services. This results in a large number of HTTP(S) requests to different services when browsing on the Web and consequently in complex traffic patterns [1], which make it difficult to detect and investigate a perpetrator's activities. In order to support investigators, we developed a methodology for visualizing a client's Web activity at a glance [2]. Our approach – Hviz – allows an investigator to quickly understand the *context* of a security alert. This can considerably speed up the bootstrapping of a forensic investigation because an analyst often starts an investigation with only very limited information (such as an IDS alert pointing to a supposedly infected client). Hviz' main focus is on visualizing the timeline of a client's Web browsing, that is, to reconstruct the sequence of Web pages visited by a user from network traces. The core idea behind Hviz is to highly aggregate regular Web activities such that anomalies clearly stand out. But recent analysis [3] shows an ongoing change in the way how bot masters control their malware remotely (C&C channel). Instead of compromising and using unpopular Web sites as C&C server, they increasingly control their malware by abusing popular Web sites. As the abused sites are completely benign and not infected, the C&C channel cannot be blocked using regular blacklists, even after discovering it. In particular, social network sites are an effective C&C channel. Using a social network as control channel can be a simple two step process: (i) the bot master posts malware instructions in a comment on a message board, (ii) the malware visits the message board from the compromised machine, downloads the corresponding Web site, and extracts the instructions. Malware analysts discovered such communication patterns for Twitter [4], Youtube and Google Plus [5], as well as Pinterest [6]. As popular, benign Web sites are used for the communication, such activities become more difficult to spot in Hviz.

Thesis goals

The aim of this thesis is to investigate and develop additional methods for interactive filtering and highlighting of network activities in Hviz. That is, to complement the already developed spatial correlation by anomaly scoring functions that incorporate the temporal sequence of events. As an example for an anomaly scoring function, one could measure the number of HTTP requests occurring while (or shortly after) an average visit to Youtube; in many cases, there will be a request loading the main page, some additional requests loading embedded objects, and a large data stream delivering the actual video. On the other hand, if a malware only fetches the main page to extract commands, these auxiliary requests will be missing. Thus, the absence of auxiliary requests can signal an anomaly.

In summary, this thesis consists of the following tasks:

- Background research on Web-based malware (e.g. [7, 8, 9, 10]).
- Experiments with real malware in a testbed to collect network traces of malicious activities.
- Analysis of typical sequences of events occurring during human Web browsing.
- Comparison of user and malware Web activity patterns.
- Development and evaluation of new anomaly scoring functions for interactive filtering of Web events.
- Write a report and present the work.

More information

For more information on this thesis, please contact David Gugelmann (gugelmann@tik.ee.ethz.ch).

References

[1] M. Butkiewicz, H. V. Madhyastha, and V. Sekar, "Understanding website complexity: Measurements, metrics, and implications," in *Proc. IMC '11*, 2011, pp. 313–328. [Online]. Available: http://doi.acm.org/10.1145/2068816.2068846

[2] D. Gugelmann, F. Gasser, B. Ager, and V. Lenders, "Hviz: Http(s) traffic aggregation and visualization for network forensics," *Digital Investigation*, vol. 12, Supplement 1, pp. 1–11, 2015, Proc. 2nd Annual DFRWS Europe (DFRWS 2015 Europe).

[3] D. Sancho, "Steganography and Malware: Concealing Code and C&C Traffic," Trend Micro – 2015-05-11, http://blog.trendmicro.com/trendlabs-security-intelligence/steganography-and-malware-concealing-code-and-cc-traffic/.

[4] "The MiniDuke Mystery: PDF 0-day Government Spy Assembler 0x29A Micro Backdoor," Kaspersky Lab – Securelist (2013-02-27), https://securelist.com/blog/incidents/31112/the-miniduke-mystery-pdf-0-day-government-spy-assembler-0x29a-micro-backdoor/.

[5] L. Franceschi-Bicchierai, "The Worst YouTube Comments Ever Were Actually Used to Control Malware," Motherboard (2015-04-22), http://motherboard.vice.com/read/the-worst-youtube-comments-ever-were-actually-used-to-control-malware.

[6] J. C. Chen, "Banking Trojan Targets South Korean Banks; Uses Pinterest as C&C Channel," Trend Micro – 2014-12-15, http://blog.trendmicro.com/trendlabs-security-intelligence/malware-campaign-targets-south-korean-banks-uses-pinterest-as-cc-channel/.

[7] N. Villeneuve and J. Bennett, "Detecting APT Activity with Network Traffic Analysis," http://www.trendmicro.com/cloud-content/us/pdfs/security-intelligence/white-papers/wp-detecting-apt-activity-with-network-traffic-analysis.pdf, Trend Micro, Tech. Rep., 2012.

[8] B. AsSadhan and J. M. F. Moura, "An efficient method to detect periodic behavior in botnet traffic by analyzing control plane traffic," *J. of Advanced Research*, vol. 5, no. 4, pp. 435 – 448, 2014, cyber Security.

[9] R. Perdisci, D. Ariu, and G. Giacinto, "Scalable fine-grained behavioral clustering of http-based malware," *Computer Networks*, vol. 57, no. 2, pp. 487 – 500, 2013, botnet Activity: Analysis, Detection and Shutdown.

[10] P. Vadrevu, B. Rahbarinia, R. Perdisci, K. Li, and M. Antonakakis, "Measuring and detecting malware downloads in live network traffic," in *Springer Computer Security – ESORICS 2013*, 2013, pp. 556–573. [Online]. Available: http://dx.doi.org/10.1007/978-3-642-40203-6_31