Long-term anomaly detection in a distributed way

Agathe Blaise[1][2] Stefano Secci[1][3], Mathieu Bouet[2], Vania Conan[2]

Despite the use of anomaly detection tools, some world-wide attacks have not been detected until too late in the last few years. One of them, the Mirai botnet, has a pattern well-recognizable though, as it scans all IP addresses on the same ports [1]. Generally, these large-scale attacks are based on ports, either newly exploited or well-known vulnerable ones such as HTTP or Telnet. Ports can be scanned to fingerprint the target machine, to exploit known vulnerabilities or to communicate with a command-and-control (C&C) server. In most cases, all of these are meant to prepare a larger attack. New attempts on one port can be observed simultaneously in the whole Internet, but surprisingly they are not detected by traditional anomaly detectors such as MAWILab [2]. There may be three reasons for that: they work on very short-term time windows, they aggregate packets by IP addresses and ISP-scale attacks are invisible at a single network scale.

In this talk, we present an anomaly detection system that spots main changes in the usage of one port. We build a long-term profile for each port and we detect anomalies via a Z-score statistical measure [3]. To distinguish between localized variations on one port and distributed significant changes, we benefit from a distributed architecture. Several monitors are sparse into the whole network so that one monitor is situated into each sub-network. They run an anomaly detection module and send the Indicators of Compromise to a central controller able to aggregate the alarms and induce an attack. We evaluated our algorithm on real traffic data and we present the results on main attacks arisen in the last three years. We also show that the distributed architecture significantly reduces the number of false positives.

References

[1] M. Antonakakis et. al, 2017, "Understanding the mirai botnet", in Proceedings of the 26th USENIX Security Symposium (USENIX Security '17).

[2] R. Fontugne et al., 2010, "MAWILab: Combining Diverse Anomaly Detectors for Automated Anomaly Labeling and Performance Benchmarking", in ACM CoNEXT.

[3] S. Seo, 2006, A Review and Comparison of Methods for Detecting Outliers in Univariate Data Sets. Masters Thesis, University of Pittsburgh.

[1] Sorbonne Université, CNRS LIP6, Paris, France, Email:{firstname.lastname}@lip6.fr
[2] Thales SIX, Gennevilliers, France. Email: {firstname.lastname}@thalesgroup.com
[3] CNAM Paris, France, Email: secci@cnam.fr.