Factors for Differentiating Human from Automated Attacks

Kelly Greeling, Graduate Student at School of Information Science - UIUC

Alex Withers, Senior Security Engineer, National Center for Supercomputing Applications - UIUC

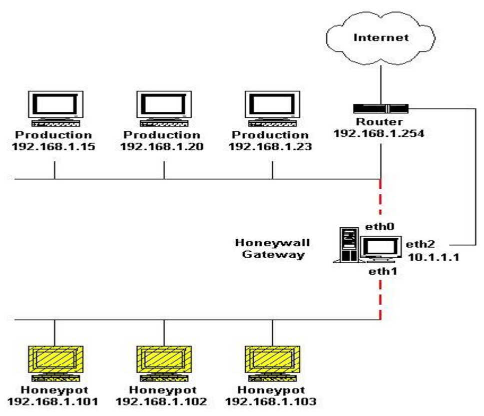
Masooda Bashir, Assistant Professor, School of Information Sciences - UIUC

Background

- Recent cyber-crime costs are at an all-time high and still skyrocketing.
- Many Intrusion Detection Systems and Intrusion Protection Systems utilize behavior-based methodology, which seeks to identify a baseline for normal users that is then used to compare against *real-time* and *non-real-time* events in an effort to locate malicious activity
- The rise of automated attacks has created a great deal of noise for security personnel to wade through to identify malicious behavior and even with IDS systems, a human actor is still required to go through the logs to note is unusual activity is actually a threat.
- If a human based attack is significantly different than an automated attack it would be extremely useful for security personnel to have a way to separate the behavior of an automated cyberattack tool from that of a human actor, as this would allow them to create separate tools to deal with each.

Research Goals

- **Long-term Project Goals**
  - Evaluate the viability of event time-difference and event pattern-occurrence as factors in behavior-based Intrusion Detection Systems for differentiating between human and automated program behavior.
  - In the future, determine how these factors can be added into Intrusion Detection Systems to help identify attackers swiftly.
- **Short-term Goals**
  - Develop and finalize protocol for capture and analysis of honeypot machine-log data administered over by the National Center for Supercomputing Applications
    - Honeypots are a type of security architecture set up to gather information on malicious activity



  - Identify any trends or regular activity in the data

Methods

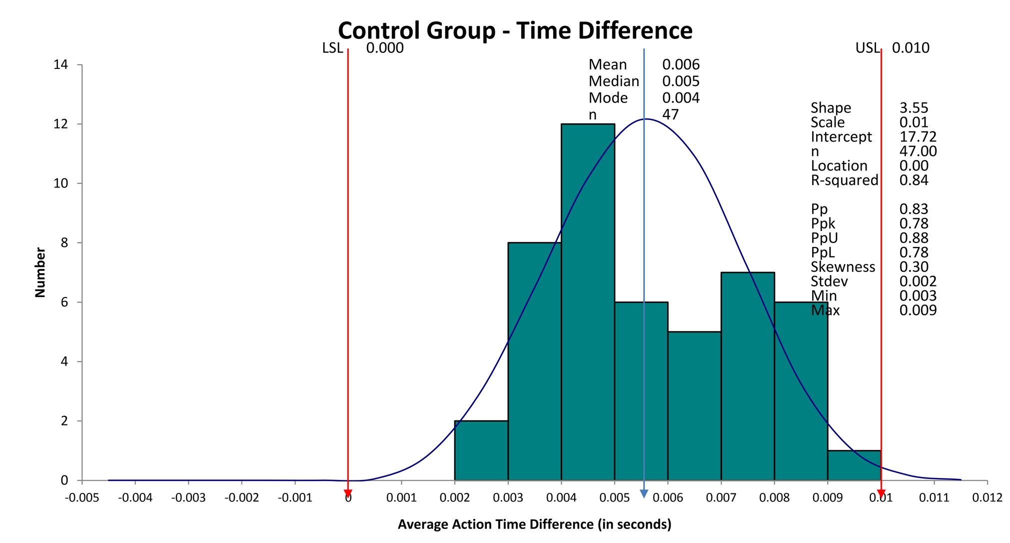
1. Organized honeypot log files by time/event/datatype

Timestamp (date)	Timestamp (time)	Event Time Difference (in seconds)	SSHD Number	Event
2015-11-12	11:42:24.172574		19698	Accepted password for … from … port … ssh2
2015-11-12	11:42:24.174622	0.002	19698	pam_unix(sshd:session): session opened for user … by (uid=0)
2015-11-12	11:42:24.981013	0.806	19698	Received disconnect from …: 11: disconnected by user
2015-11-12	11:42:24.981024	0	19698	pam_unix(sshd:session): session closed for user …

2. Employed Syntactic Pattern Recognition of events in order to establish patterns

Timestamp (date)	Timestamp (time)	Event Time Difference (in seconds)	SSHD	SSHD Pattern
2015-11-12	11:42:24.172574		19698	
2015-11-12	11:42:24.174622	0.002	19698	
2015-11-12	11:42:24.981013	0.806	19698	
2015-11-12	11:42:24.981024	0	19698	7 8 5 9

3. Pulled CRON (known program) patterns/times/frequency to form control



Preliminary Results

- Protocol creation complete
- **Trends and Regular Activity**
- The entire network test group (n=63) averaged 4.67 ± 1.88.
- The combined keystroke test group (n=190) averaged .26 ± .04 seconds.
- The keystroke data revealed four unknown *Pattern Groups*, two of which were individual events.
- While the network group had a total of seven unknown *Pattern_Groups,* two of which were individual event occurrences.
- In both the keystroke and network test groups there were several *Pattern_Groups* that occurred very quickly within a small duration of time. (see Figure 1
- There were also groups that took significantly longer to occur and were rarer. (see Figure 2)

Figure 1:

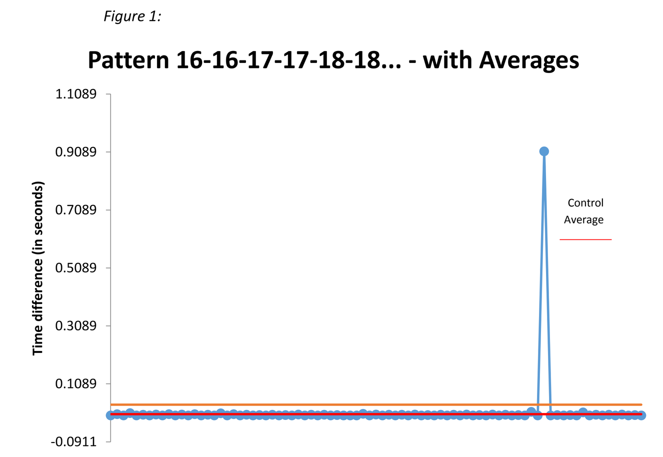
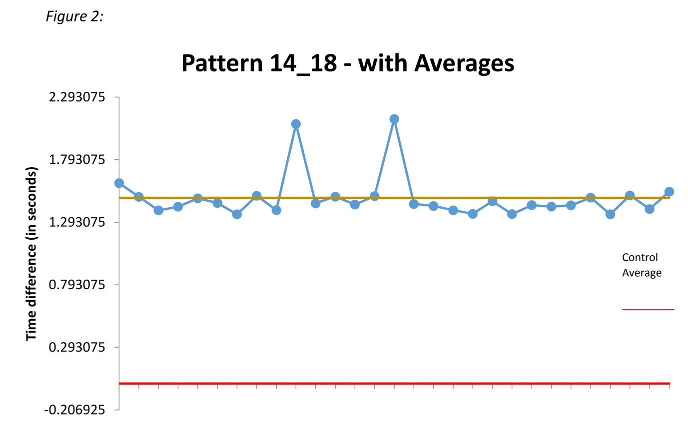


Figure 2:



Conclusions

- Some groups complete events within a rapid period of time, and repeat the same pattern of events over and over with little to no deviation.
- Other groups take a longer period of time to complete events and fall outside the standard deviation.
- This initial research has shown that *Pattern_Occurrence* and *Time_Difference* are indeed likely viable factors to separate human behavior from automated program behavior in an IDS and need further study

Future Research

- Obtain larger sample size to replicate preliminary results and improve statistical signifigance
- Establishing a way to add normalized human behavior data (as honeypots servers, by design, do not have regular users)
- Designing an experiment to control for issues like distance-from-server lag, IP bounce, etc.

Acknowledgments

This project would not have been possible without help from my advisor, Dr. Masooda Bashir or assistance and data access from Alex Withers and the NCSA. Nor, without technical assistance from Bartosz G. Kosciarz and Seoung Kyun Kim. This material is based upon work supported by the Maryland Procurement Office under Contract No. H98230-14-C-0141

References

- M. Rogers, "The role of criminal profiling in the computer forensics process," *Comput. Secur.*, vol. 22, no. 4, pp. 292–298, 2003.
- M. K. Rogers, "Psychological profiling as an investigative tool for digital forensics," *Digit. Forensics Threat. Best Pract.*, p. 45, 2015.
- M. K. Rogers and K. Seigfried, "The future of computer forensics: a needs analysis survey," *Comput. Secur.*, vol. 23, no. 1, pp. 12–16, 2004.
- A. Reyes, K. O'Shea, J. Steele, J. R. Hansen, B. R. Jean, and T. Ralph, "Chapter 2 - 'Computer Crime' Discussed," in *Cyber Crime Investigations*, Burlington: Syngress, 2007, pp. 23–47.
- L. Kwan, P. Ray, and G. Stephens, "Towards a methodology for profiling cyber criminals," presented at the Hawaii International Conference on System Sciences, Proceedings of the 41st Annual, 2008, pp. 264–264.
- C. M. Colombini and A. Colella, "Digital scene of crime: technique of profiling users.," *JoWUA*, vol. 3, no. 3, pp. 50–73, 2012.
- C. Colombini and A. Colella, "Digital Profiling: A Computer Forensics Approach," in *Availability, Reliability and Security for Business, Enterprise and Health Information Systems: IFIP WG 8.4/8.9 International Cross Domain Conference and Workshop, ARES 2011, Vienna, Austria, August 22-26, 2011. Proceedings*, A. M. Tjoa, G. Quirchmayr, I. You, and L. Xu, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2011, pp. 330–343.
- E. V. Linden, *Focus on terrorism*, vol. 9. Nova Publishers, 2007.
- "About time ranges in search - Splunk Knowledgebase." [Online]. Available: http://docs.splunk.com/Documentation/Splunk/6.1.7/Search/Aboutsearchtimeranges. [Accessed: 22-Jul-2016].
- J. D. Brutlag, "Aberrant Behavior Detection in Time Series for Network Monitoring.," presented at the LISA, 2000, vol. 14, pp. 139–146.
- L. Spitzner, "Know your enemy: Honeynets, what a honeynet is, its value, overview of how it works, and risk/issues involved," *Web May*, 2006.
- Ponemon Institute, "2015 cost of cyber crime study: Global," Hewlett Packer Enterprise, Oct. 2015.

HOTSOS in the SCIENCE OF SECURITY SYMPOSIUM & BOOTCAMP

HoTSoS Symposium and Bootcamp

HOT TOPICS *in the* **SCIENCE OF SECURITY**

APRIL 4-5, 2017 | HANOVER, MARYLAND