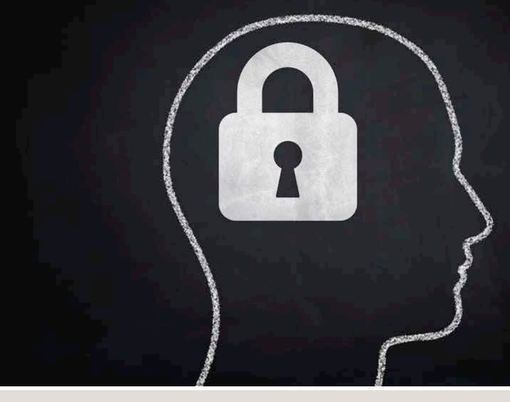
INTRUSION DETECTION:

SEPARATING THE HUMAN FROM THE PROGRAM

Kelly Greeling, Alex Withers, Masooda Bashir



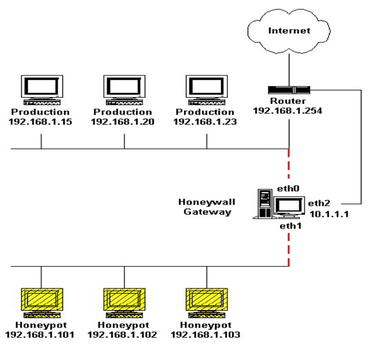
Problem 1: The rise of automated attacks has created a great deal of noise for security personnel to wade through to identify malicious behavior

Problem 2: While many real-time detection tools and log management programs consider time in terms of the repetition of events over a time series, attackers have caught onto this and begun varying the time between their attacks

Research Goal: This project utilized syntactic pattern recognition techniques combined with an analysis of event time difference to discover if there is a statistical way in which to separate human from automated intrusion behavior in a system

Methods

- Honeypots are a type of security architecture set up to gather information on malicious activity



  - Any activity with the honeypot is taken as malicious, as the honeypot runs no authorized services
  - Therefore, honeypots make a perfect testbed for cybersecurity researchers

Adapted (and image taken) from: L. Spitzner, "Know your enemy: Honeynets, what a honeynet is, its value, overview of how it works, and risk/issues involved," *Web May*, 2006.

- Pulled Auth and Snoopy logs from a set of honeypots administered over by the NCSA

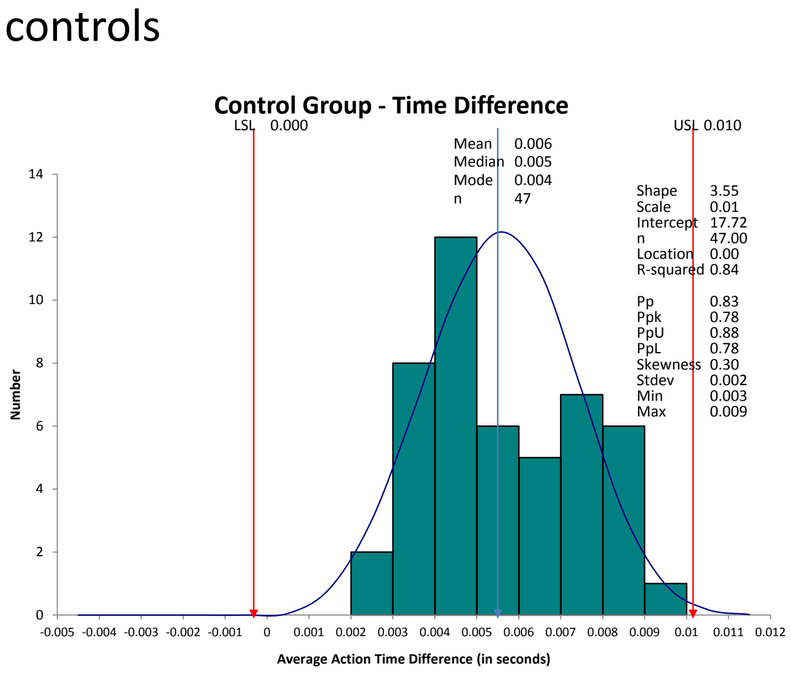
- Organized honeypot log files by time/event/datatype (SID, SSHD, CRON, etc.)

Timestamp (date)	Timestamp (time)	Event Time Difference (in seconds)	SSHD Number	Event
2015-11-12	11:42:24.172574		19698	Accepted password for root from 130.126.137.23 port 51881 ssh2
2015-11-12	11:42:24.174622	0.002	19698	pam_unix(sshd:session): session opened for user root by (uid=0)
2015-11-12	11:42:24.981013	0.806	19698	Received disconnect from 130.126.137.23: 11: disconnected by user
2015-11-12	11:42:24.981024	0	19698	pam_unix(sshd:session): session closed for user root

- 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

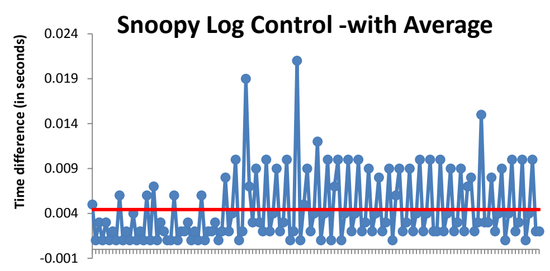
- Pulled CRON (known program) patterns/times/frequency to form controls

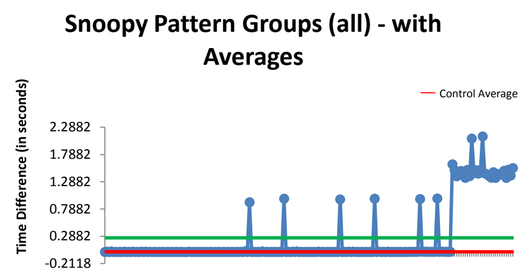


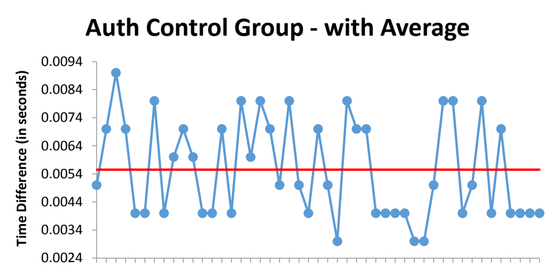
- Compared each pattern-group against controls

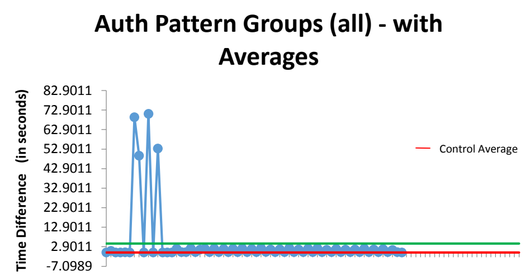
Results

- When comparing event time difference of the pattern-groups (all) against the control, there are several notable events that are statistically different

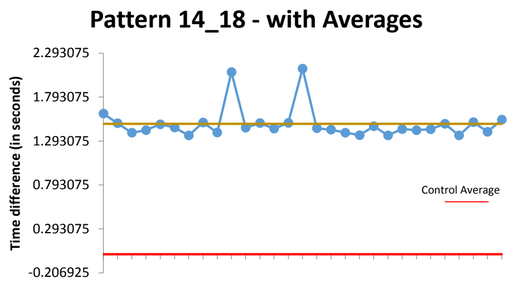


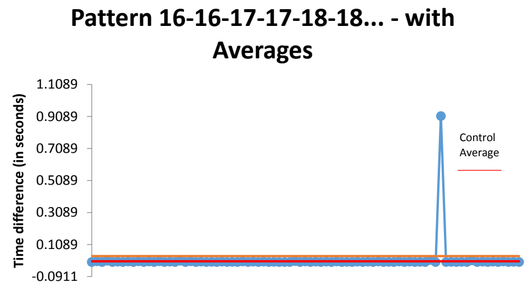






- This becomes even more apparent when you compare the individual pattern-groups against the control mean





Conclusions and Future Research

- More research is needed to find other variables to control for within time difference. (e.g. connection reliability and speed, IP spoofing, etc.)
- More data should be collected to test, in order to conduct a more detailed analysis
- Conduct a more detailed linguistic breakdown of the patterns themselves
- **It seems clear, however, that investigating event time difference between events as an indicator for separating human from automated program behavior in a system should be made a priority**

Acknowledgements

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.

SCIENCE OF SECURITY
VIRTUAL ORGANIZATION
Funded by the National Security Agency.

INFORMATIONTRUST INSTITUTE