APPLYING SECURITY ORCHESTRATION, AUTOMATION, & RESPONSE (SOAR) TO SECURITY THREAT HUNTING WITH DUKE'S IT SECURITY OFFICE

PROJECT LEADS: PHILLIP BATTON, ERIC HOPE

PROJECT MANAGER: JOAO MANSUR

TEAM: MATT FEDER, VARUN PRASAD, JOHN TAYLOR

OVERVIEW/BACKGROUND

Duke is a prime target for cybercriminals

Proactive approach to security threat hunting

Exploratory analysis of web log data from May 15, 2020

PROJECT GOALS

FIND UNIQUE/ANOMALOUS REQUESTS IN LOGS

ANALYZE TRENDS/PATTERNS IN DATA

IDENTIFY REQUESTS THAT ARE POTENTIALLY MALICIOUS TO HELP OIT

WEB LOG DATA

- May 15, 2020
  - Excludes campus IPs and authenticated requests
- Key variables: host, hashed ip, time, query, useragent

	host	source	clientip	ident	user	_time	uri_query	status	bytes	referer	useragent	clientip_hash
0	fleet.oit.duke.edu	/var/log/httpd/ssl_access_log.4.gz	xxx.xxx	-	-	2020-05-15T15:59:59.000-0400	NaN	200	20	-	osquery/2.10.2	
1	sites.duke.edu	/var/log/httpd/ssl_access_log.1	xxx.xxx	-	-	2020-05-15T15:59:59.000-0400	target=https%3A%2F%2Fsites.duke.edu%2Fwp-login...	302	741	-	UT-Dorkbot/1.0	
2	sites.duke.edu	/var/log/httpd/ssl_access_log.2.gz	xxx.xxx	-	-	2020-05-15T15:59:59.000-0400	target=https%3A%2F%2Fsites.duke.edu%2Fwp-login...	302	741	-	UT-Dorkbot/1.0	
3	fleet.oit.duke.edu	/var/log/httpd/ssl_access_log.4.gz	xxx.xxx	-	-	2020-05-15T15:59:59.000-0400	NaN	200	20	-	osquery/2.10.2	
4	fleet.oit.duke.edu	/var/log/httpd/ssl_access_log.2.gz	xxx.xxx	-	-	2020-05-15T15:59:59.000-0400	NaN	200	20	-	osquery/2.10.2	

BADACTORS CACHE

- File of blocked IPs from May 15, 2020
- Hashed ip, attack description, attack count
- Used for matching with web data and for data analysis

Description <chr>	n <int>
chn honeypot detected scanner	29448
netflow detected scanner	7146
IPS IDed SSH brute force useragents	930
WP brute force attempts in 24 hours	175
IPS IDed suspicious file attempts	152
IPS IDed mirai and reaper attacks	109
blocking slash28 due to percentage already blocked	94
Sipvicious	38
ssh invalid users in 24 hours	36
masscan attempts	33

PROCESS

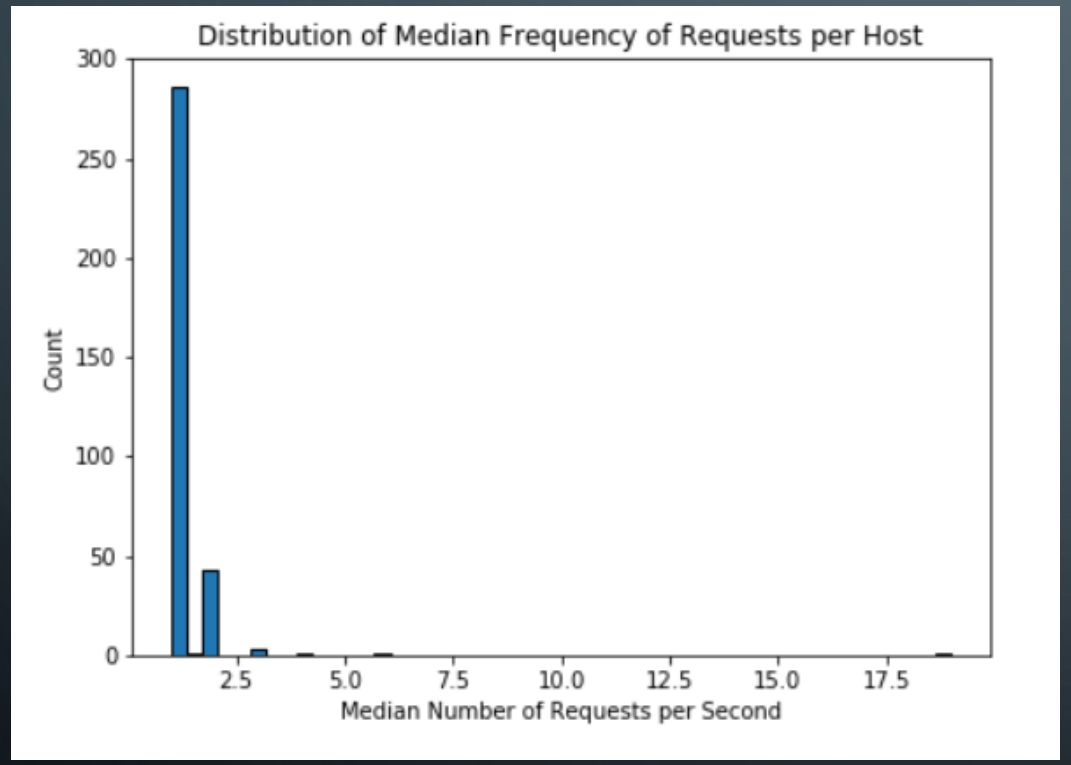
MERGE WEB LOG DATA WITH BADACTORS CACHE

ANALYZE PATTERNS ACROSS DIFFERENT HOSTS AND IPS

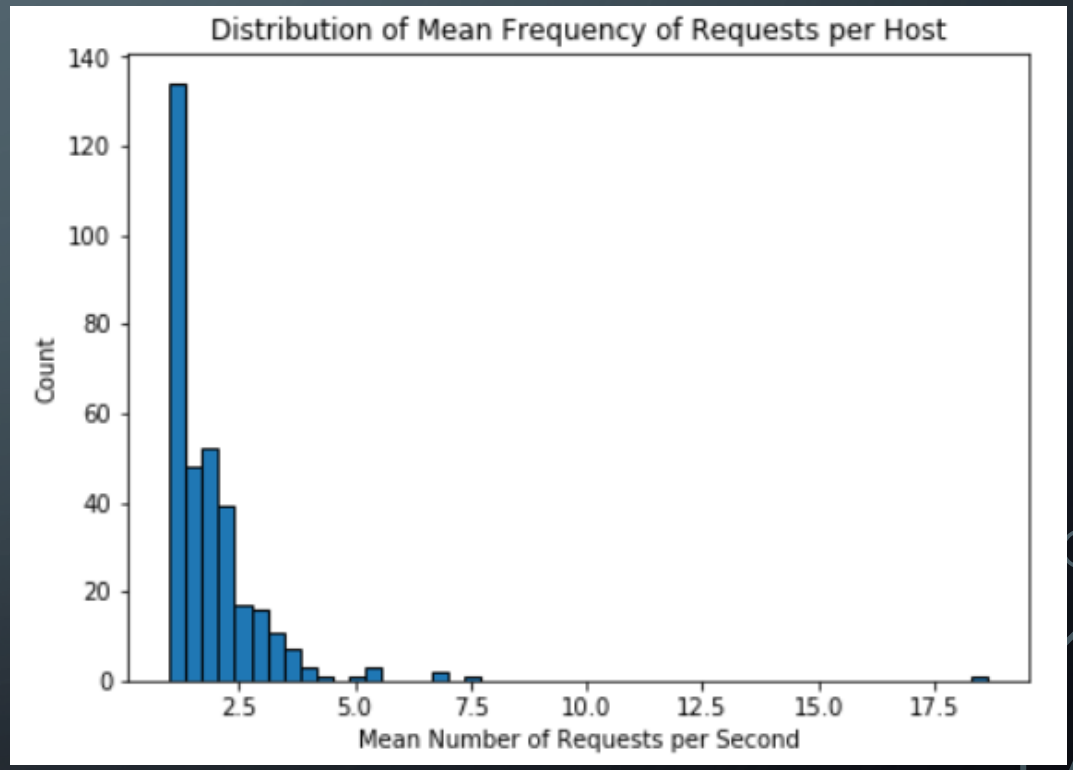
TIME SERIES ANALYSIS ON REQUESTS MADE BY NORMAL AND BAD ACTORS

MAJORITY OF HOSTS HAVE LESS THAN 5 REQUESTS PER SECOND ON AVERAGE

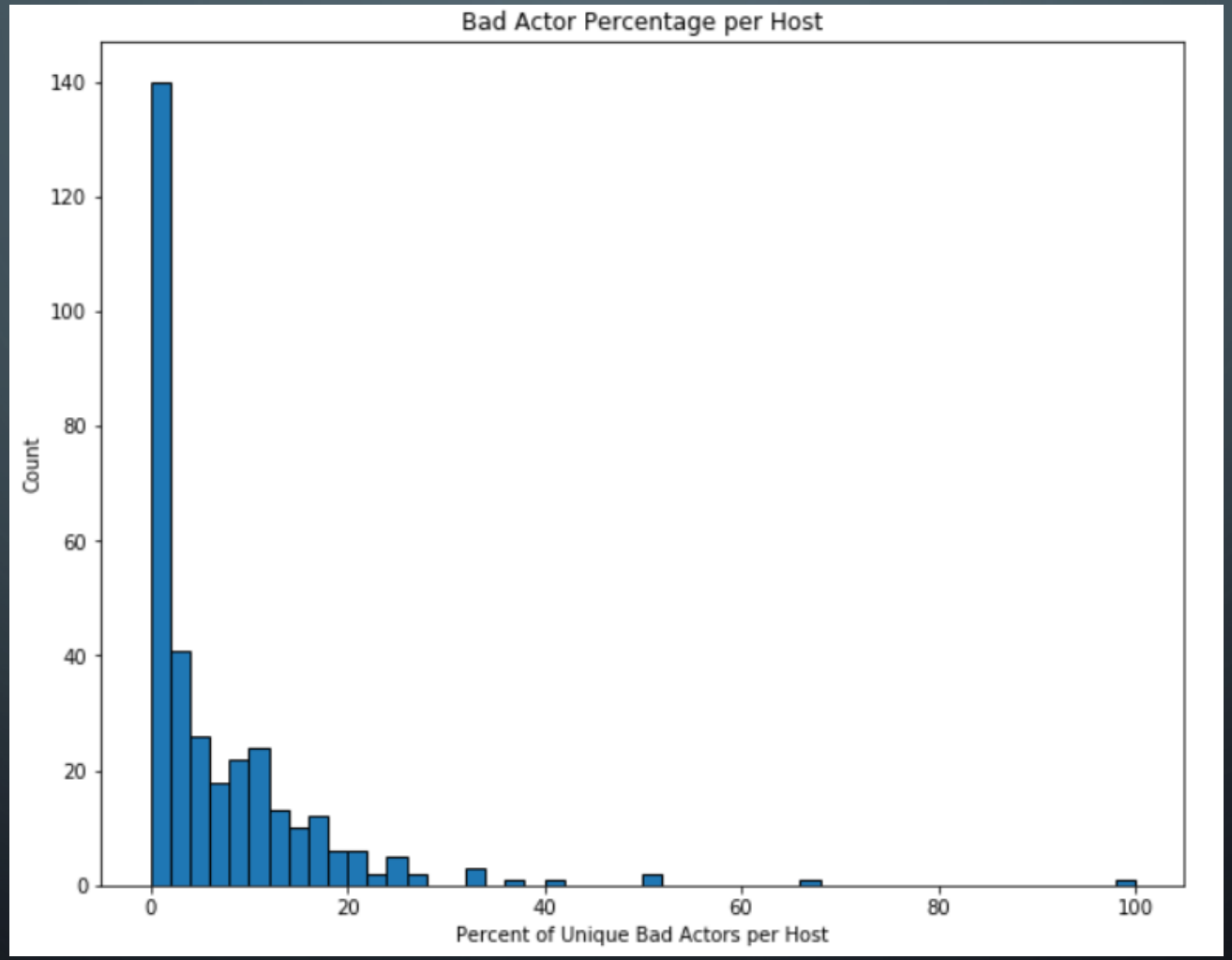
MEDIAN



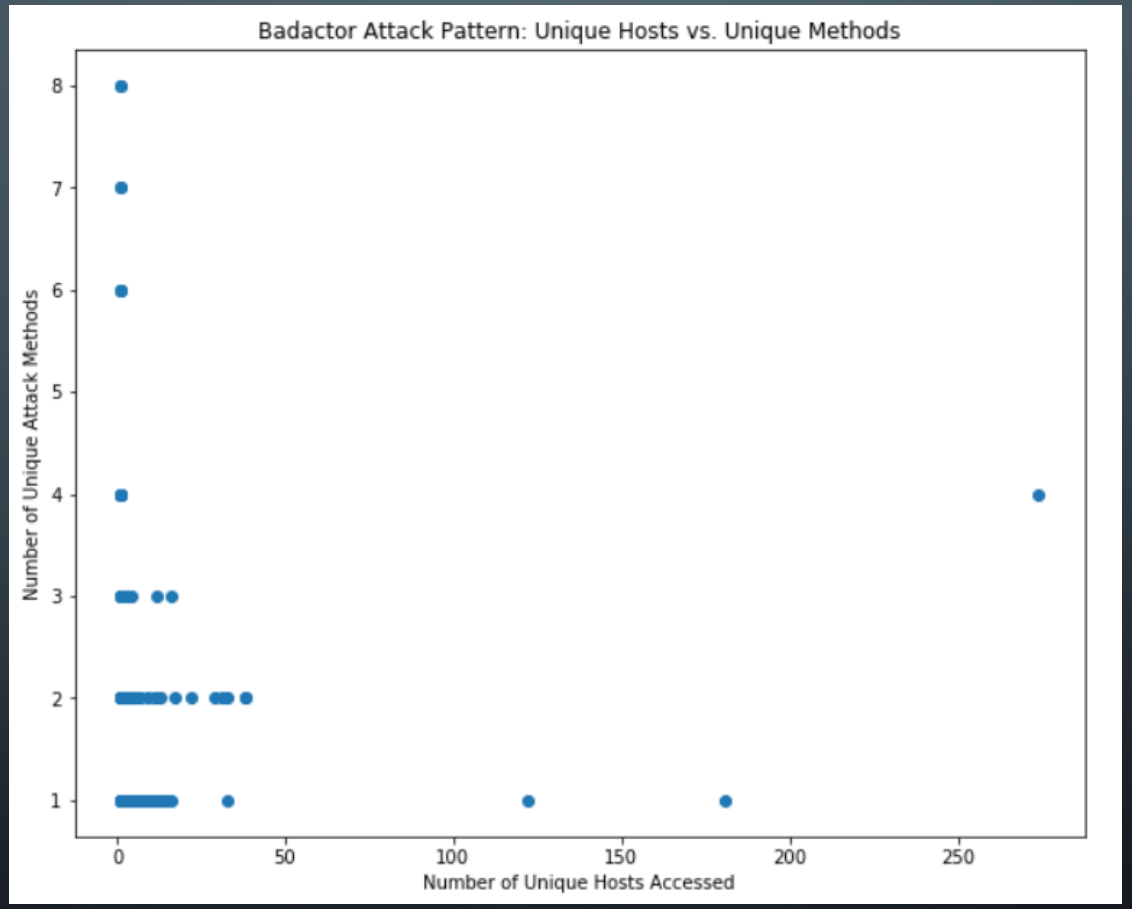
MEAN



DISTRIBUTION OF PERCENT OF BADACTORS PER HOST

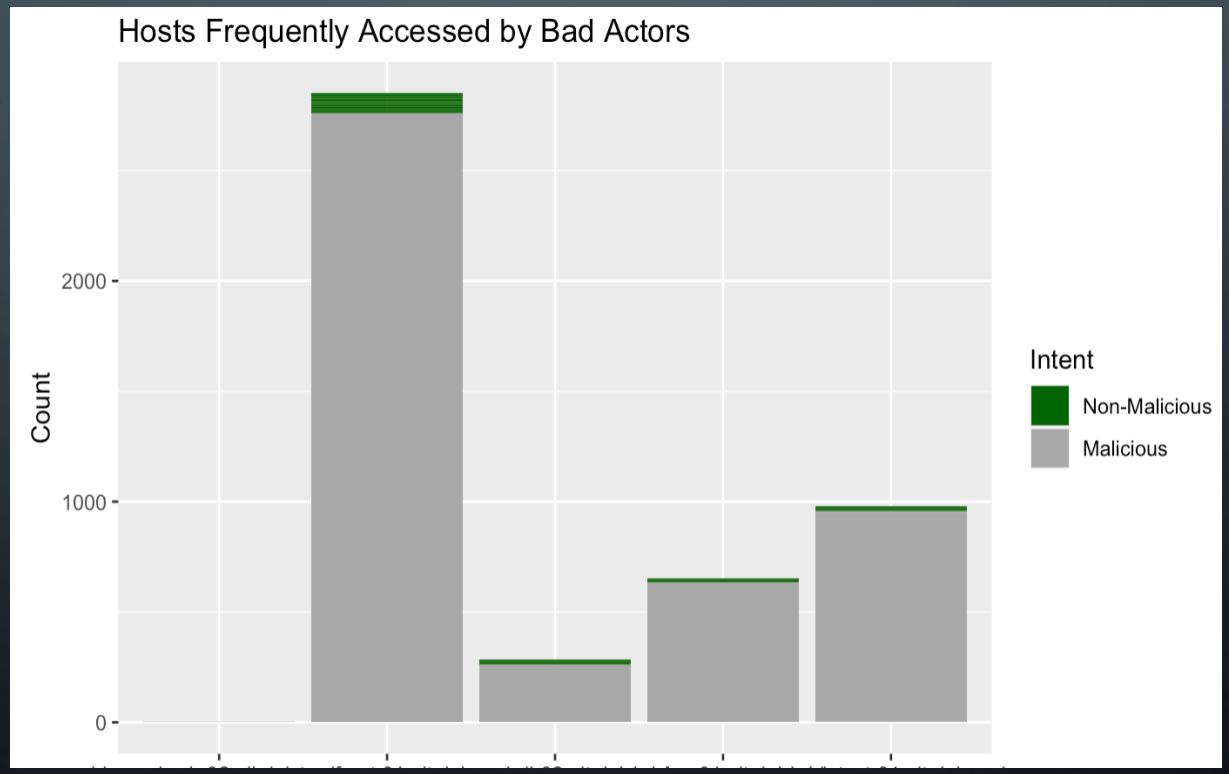


BAD ACTORS THAT USE MULTIPLE ATTACKS TEND TO FOCUS ON ONE HOST



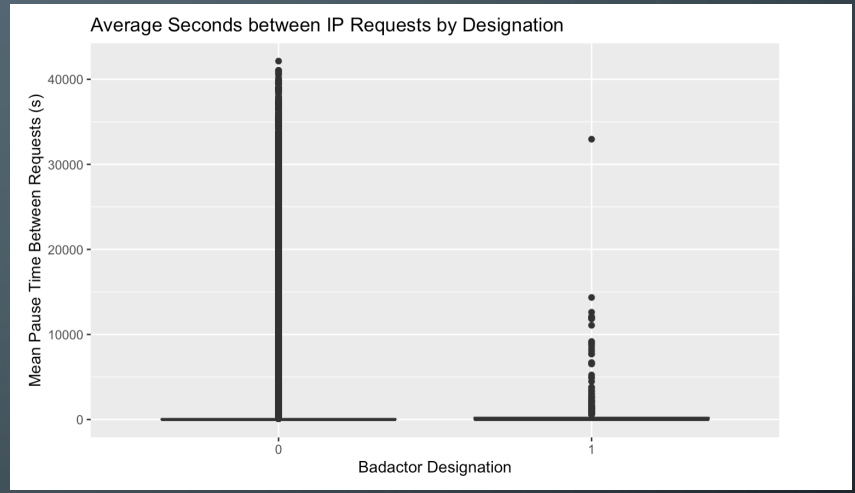
SEVERAL HOSTS HAVE OVER 90% OF REQUESTS FROM BAD ACTORS

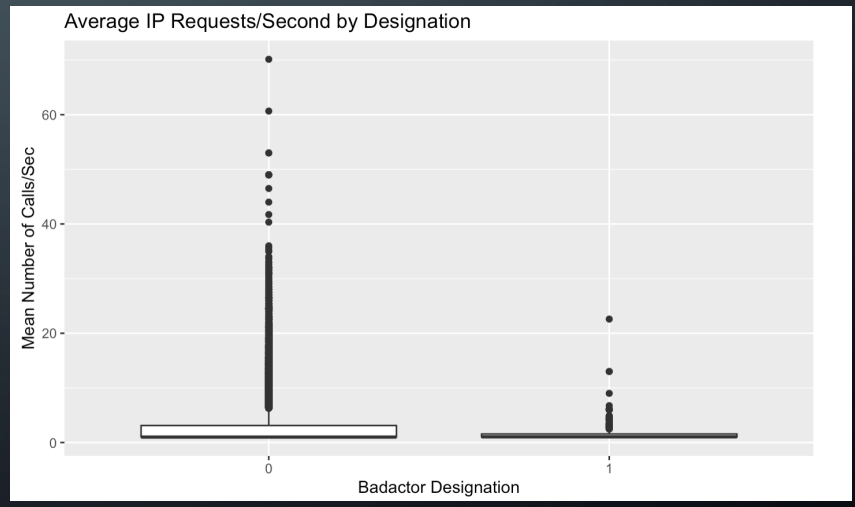
	idms-sdweb-02.oit.duke.edu	**law-storefront-01.oit.duke.edu***	mass-email-02.oit.duke.edu	web-ichw-01.oit.duke.edu	web-irb-test-01.duke.edu
Badactors	2	2754	259	633	952
Normal Actors	0	88	1	6	6



NORMAL AND BAD ACTORS SHOW DIFFERENT DISTRIBUTIONS IN CALL PATTERNS

- Finding unusual call counts/patterns
  - What qualifies as unusual?
- Distributions differ
  - More observations in the “non-badactor” group
  - More skewness





ACCOMPLISHMENTS

2 new detection techniques already operationalized from our results

New signatures added based on our analysis of ip request intent ratios

OIT has made a to-do list of other signatures to add to detection systems

Based on time, host proportion, and, method frequency

FUTURE WORK

- Flagging hosts with large proportions of “badactor” access for additional screening
  - Law-storefront still getting repeated attacks
- Analyze IP patterns within each host of interest
  - Further classify bot vs. human IPs
- Utilize the differences in variables for good/bad requests
  - Create a model with these variables
  - Contribute to a “risk score” system
- Real-time ML model

ACKNOWLEDGEMENTS

Eric Hope and Phillip Batton

Joao Mansur

Duke OIT

Cisco

Data+

THANK YOU FOR LISTENING!

ANY QUESTIONS?