

Enabling Network Visibility and Security through Tensor Analysis

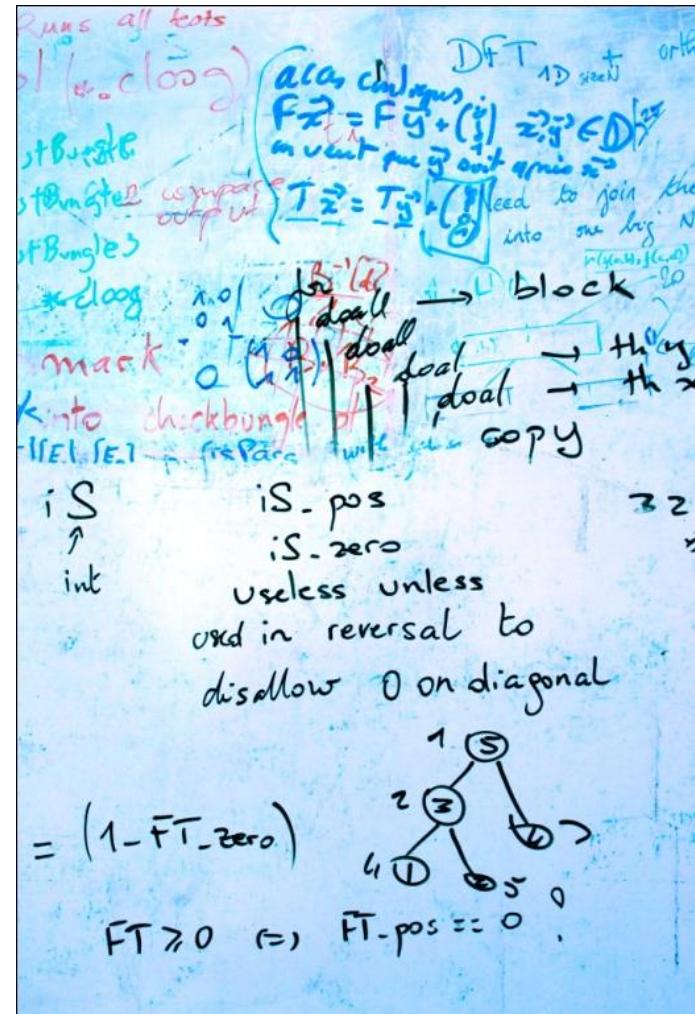
INDIS 2017

Muthu Baskaran
David Bruns-Smith[#]
Thomas Henretty
James Ezick
Richard Lethin

Reservoir Labs

[#] UC Berkeley (Work done at Reservoir Labs)

12 November 2017



www.reservoir.com

Value Add to the Security Ecosystem

We have tools that excel at monitoring flows

- Signature-based tools that provide real-time alerts
 - More advanced metadata collection tools enable deeper offline analysis

Challenges

- Rules can be anticipated and evaded
 - Metadata analysis is a forensic activity
 - Both require a known starting point

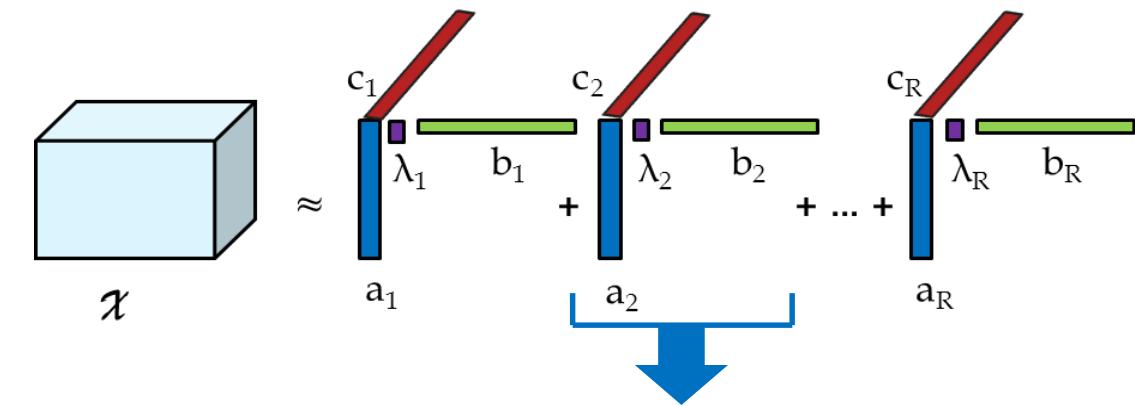
Question

- Can we supplement traditional incident-focused approaches to threat discovery with an approach that feeds metadata to a pattern-focused analytic?

01/15/2015-18:17:24.911691 [**] [1:2018635:8] ET TROJAN Common Upatre Header Structure 2 [**] [C]
A Network Trojan was detected [Priority: 1] {TCP} 172.16.120.154:49411 -> 66.147.240.173:80
01/15/2015-18:15:59.334121 [**] [1:2020212:6] ET CURRENT EVENTS Upatre Redirector IE Requesting F
69.89.27.218:80
01/15/2015-18:15:58.416495 [**] [1:2020304:2] ET CURRENT_EVENTS Upatre Redirector Jan 23 2015 [**]
[Classification: A Network Trojan was detected] [Priority: 1] {TCP} 172.16.120.154:49378 -> 69.89.
01/15/2015-18:15:58.708129 [**] [1:2020159:6] ET CURRENT_EVENTS Upatre Redirector Jan 9 2015 [**]
[Classification: A Network Trojan was detected] [Priority: 1] {TCP} 172.16.120.154:49378 -> 69.89.
01/15/2015-18:17:29.752647 [**] [1:2020235:3] ET TROJAN Mozilla Suspicious User-Agent Jan 15 2015
[Classification: A Network Trojan was detected] [Priority: 1] {TCP} 172.16.120.154:49413 -> 202.15
01/15/2015-18:17:29.752647 [**] [1:2018369:3] ET TROJAN Common Upatre URI/Headers Struct [**] [C]
A Network Trojan was detected [Priority: 1] {TCP} 172.16.120.154:49413 -> 202.153.35.133:29103
09/18/2015-20:03:13.532117 [**] [1:2012648:3] ET POLICY Dropbox Client Broadcasting [**] [Classif
Potential Corporate Privacy Violation] [Priority: 1] {UDP} 10.211.55.2:17500 -> 10.211.55.255:1756
09/18/2015-20:10:44.276238 [**] [1:2012648:3] ET POLICY Dropbox Client Broadcasting [**] [Classif
Potential Corporate Privacy Violation] [Priority: 1] {UDP} 10.211.55.2:17500 -> 10.211.55.255:1756
09/18/2015-20:13:16.104238 [**] [1:2013504:4] ET POLICY GNU/Linux APT User-Agent Outbound likely
package management [**] [Classification: Not Suspicious Traffic] [Priority: 3] {TCP} 10.211.55.8:4
91.189.91.24:80
09/18/2015-20:13:16.245969 [**] [1:2013504:4] ET POLICY GNU/Linux APT User-Agent Outbound likely
package management [**] [Classification: Not Suspicious Traffic] [Priority: 3] {TCP} 10.211.55.8:4
91.189.91.24:80
09/18/2015-20:13:16.438833 [**] [1:2013504:4] ET POLICY GNU/Linux APT User-Agent Outbound likely

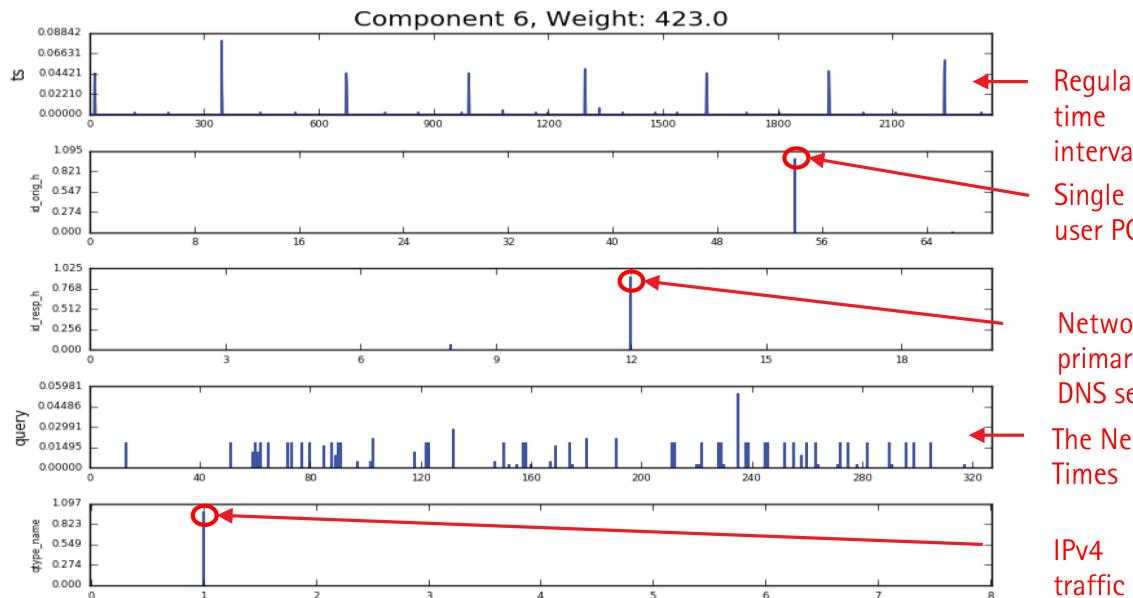
Tensor Analysis and Tensor Decompositions

A New Paradigm for Network Analysis



Unsupervised learning

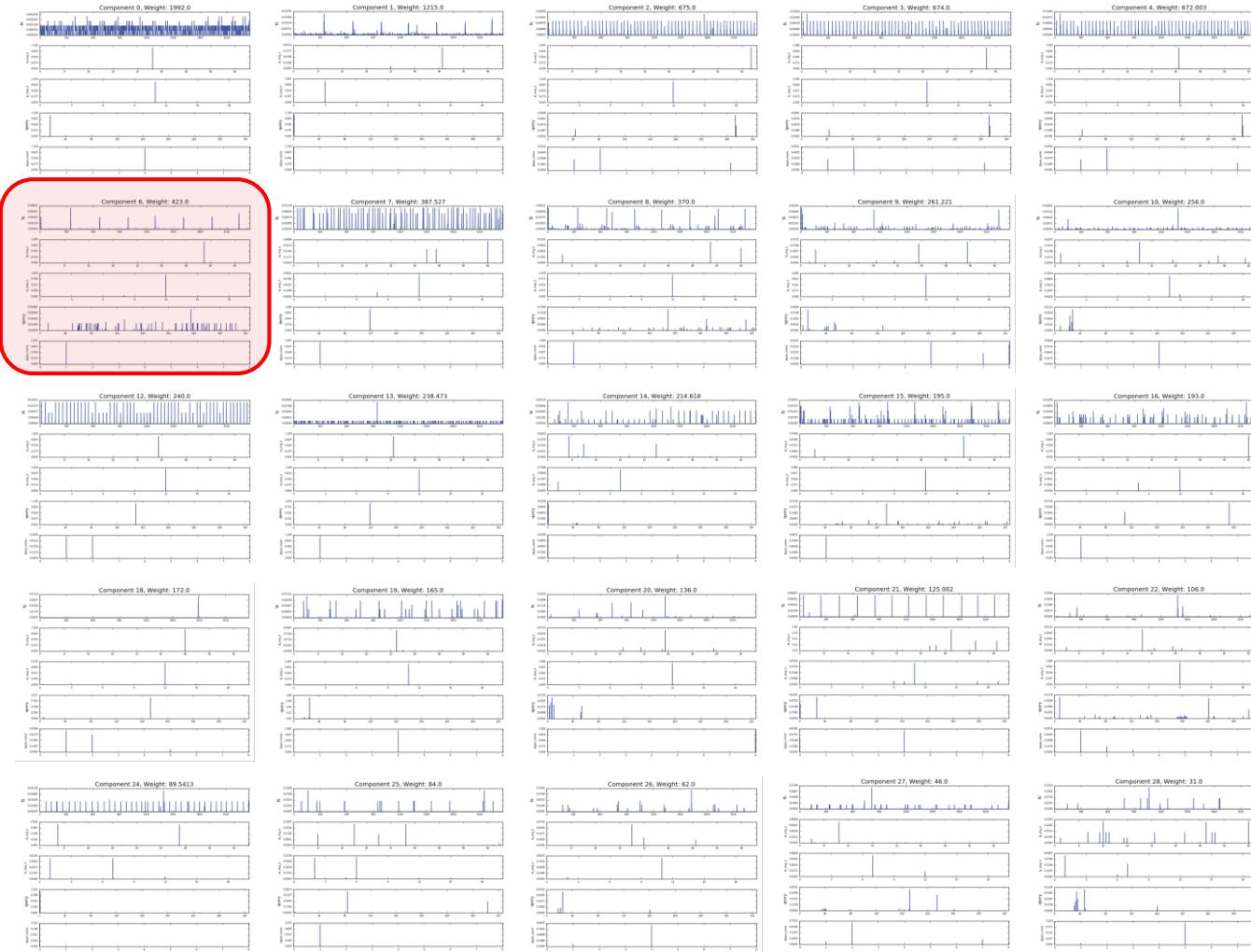
Detects unknown
unknowns



Captures coherent
patterns of activity
spanning multiple
dimensions

Tensor Analysis and Tensor Decompositions

A New Paradigm for Network Analysis



Lesser cognitive load for analyst

Looks at fewer components (indicating activities of interest)

Use the patterns of activity to guide further investigation

CANDID and ENSIGN: Context and Overview

CANDID: A tool for network security and traffic analysis

- Provides comprehensive and contextual insights into the network
 - Malicious and obfuscated network threats
 - Network state and network performance indicators
- Reduces the cognitive load of network analysts

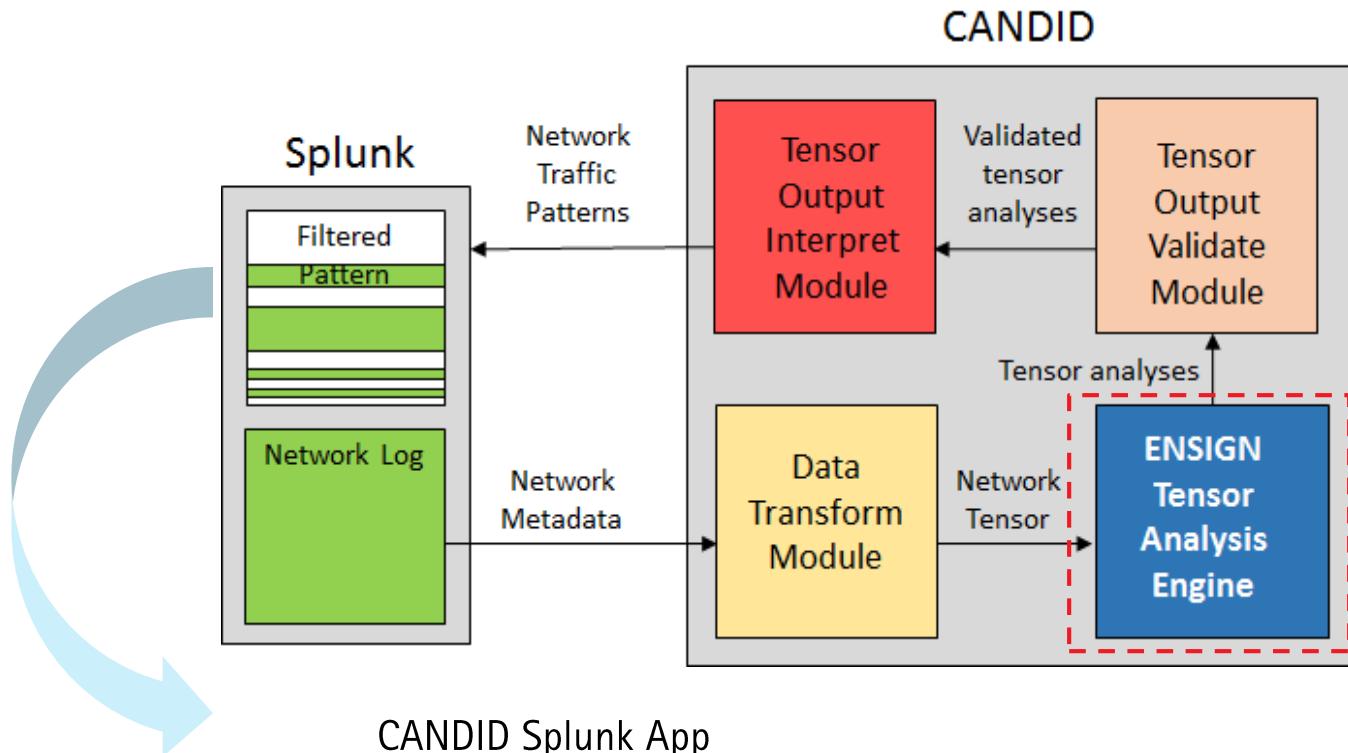
ENSIGN: High-performance tensor analysis engine driving CANDID

- Tensor Toolbox with advanced mathematical methods for data analysis
- High performance, rich capability, easy usability

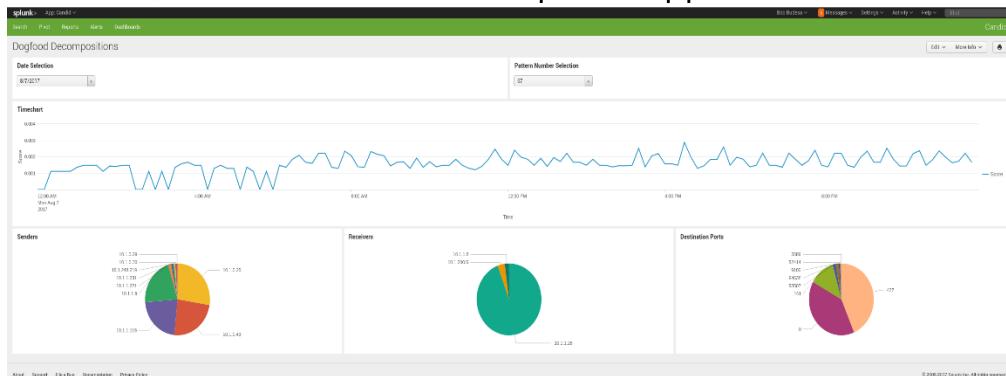
Successfully used in diverse operational environments

- Security Operations Center (SOC) for the SCinet network at SC16
- Reservoir Labs' Local Area Network (LAN)

CANDID : Tool workflow



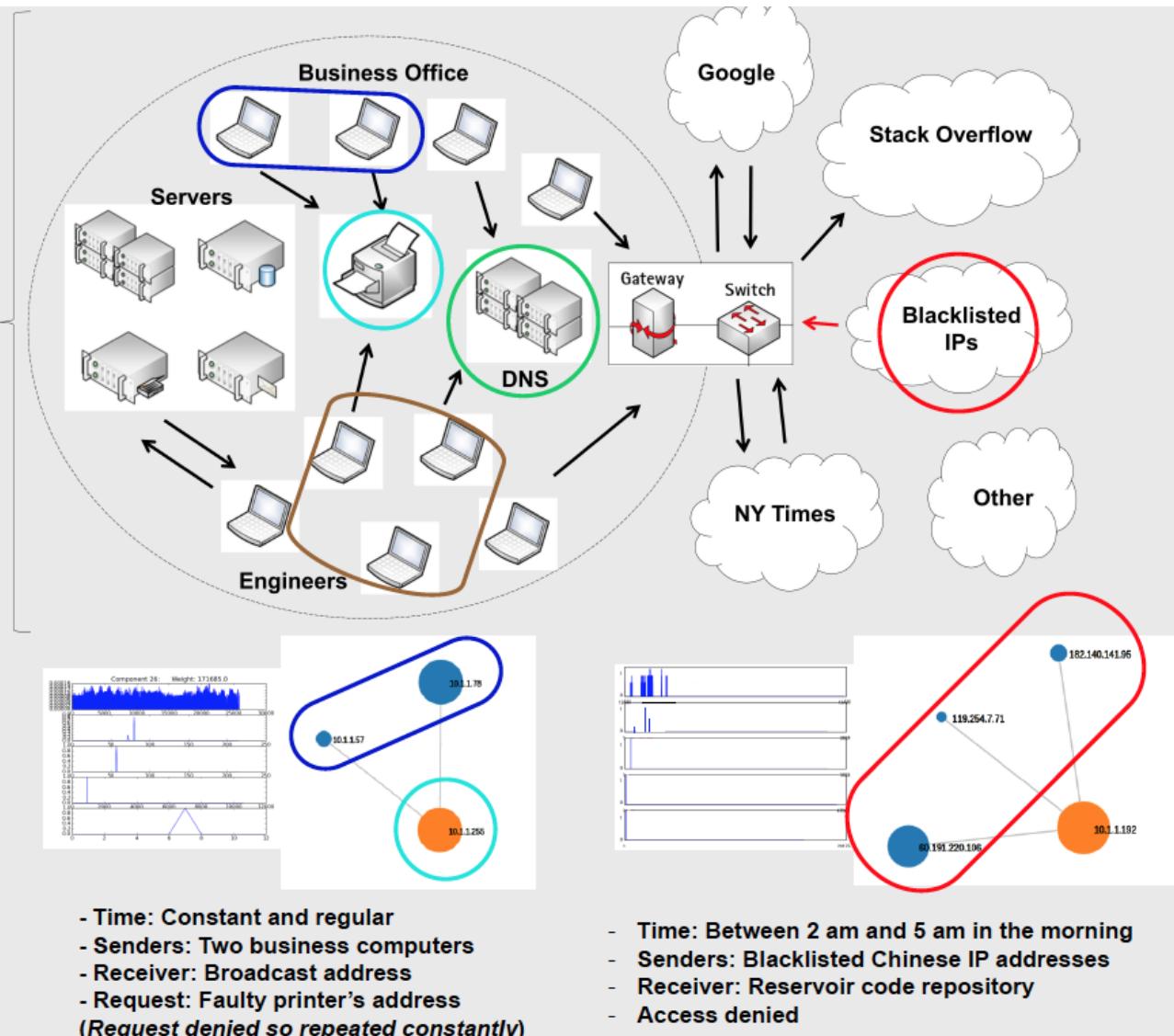
CANDID Splunk App



CANDID on Reservoir Network Traffic

**Raw Data from R-Scope®
(Reservoir's Network Security Monitoring Appliance)**

Generator	Time	string	addr	port	id.resp_h	id.resp_p	proto
140484995_10902	2017-07-27 11:00:00	cassandra1595095a	10.1.1.8	90188	64.72.64.10	53	3
140484995_128403	2017-07-27 11:00:00	cassandra1595095a	10.1.1.8	90208	64.72.64.10	53	3
140484995_130826	2017-07-27 11:00:00	cassandra1595095a	10.1.1.8	90617	64.72.64.10	53	3
140484995_130827	2017-07-27 11:00:00	cassandra1595095a	10.1.1.8	90628	64.72.64.10	53	3
140484995_130828	2017-07-27 11:00:00	cassandra1595095a	10.1.1.8	90629	64.72.64.10	53	3
140484995_130829	2017-07-27 11:00:00	cassandra1595095a	10.1.1.8	90630	64.72.64.10	53	3
140484995_242368	2017-07-27 11:00:00	cassandra934a0553	10.1.1.8	90705	64.72.64.10	53	3
140484995_439768	2017-07-27 11:00:00	cassandra934a0553	10.1.1.8	90818	64.72.64.10	53	3
140484995_439769	2017-07-27 11:00:00	cassandra934a0553	10.1.1.8	90819	64.72.64.10	53	3
140484995_439770	2017-07-27 11:00:00	cassandra934a0553	10.1.1.8	90820	64.72.64.10	53	3
140484995_510783	2017-07-27 11:00:00	cassandra934a0553	10.1.1.8	90723	64.72.64.10	53	3
140484995_512208	2017-07-27 11:00:00	cassandra934a0553	10.1.1.8	91156	64.72.64.10	53	3
140484995_512209	2017-07-27 11:00:00	cassandra934a0553	10.1.1.8	91203	64.72.64.10	53	3
140484995_152038	2017-07-27 11:00:00	cassandra934a0553	10.1.1.8	90648	64.72.64.10	53	3
140484995_154016	2017-07-27 11:00:00	cassandra934a0553	10.1.1.8	90649	64.72.64.10	53	3
140484995_154017	2017-07-27 11:00:00	cassandra934a0553	10.1.1.8	90650	64.72.64.10	53	3
140484995_180541	2017-07-27 11:00:00	cassandra934a0553	10.1.1.8	90651	64.72.64.10	53	3
140484995_180542	2017-07-27 11:00:00	cassandra934a0553	10.1.1.8	90652	64.72.64.10	53	3
140484995_180543	2017-07-27 11:00:00	cassandra934a0553	10.1.1.8	90653	64.72.64.10	53	3
140484995_180544	2017-07-27 11:00:00	cassandra934a0553	10.1.1.8	90654	64.72.64.10	53	3
140484995_180545	2017-07-27 11:00:00	cassandra934a0553	10.1.1.8	90655	64.72.64.10	53	3
140484995_180546	2017-07-27 11:00:00	cassandra934a0553	10.1.1.8	90656	64.72.64.10	53	3
140484995_180547	2017-07-27 11:00:00	cassandra934a0553	10.1.1.8	90657	64.72.64.10	53	3
140484995_180548	2017-07-27 11:00:00	cassandra934a0553	10.1.1.8	90658	64.72.64.10	53	3
140484995_180549	2017-07-27 11:00:00	cassandra934a0553	10.1.1.8	90659	64.72.64.10	53	3
140484995_180550	2017-07-27 11:00:00	cassandra934a0553	10.1.1.8	90660	64.72.64.10	53	3
140484995_180551	2017-07-27 11:00:00	cassandra934a0553	10.1.1.8	90661	64.72.64.10	53	3
140484995_180552	2017-07-27 11:00:00	cassandra934a0553	10.1.1.8	90662	64.72.64.10	53	3
140484995_180553	2017-07-27 11:00:00	cassandra934a0553	10.1.1.8	90663	64.72.64.10	53	3
140484995_180554	2017-07-27 11:00:00	cassandra934a0553	10.1.1.8	90664	64.72.64.10	53	3
140484995_180555	2017-07-27 11:00:00	cassandra934a0553	10.1.1.8	90665	64.72.64.10	53	3
140484995_180556	2017-07-27 11:00:00	cassandra934a0553	10.1.1.8	90666	64.72.64.10	53	3
140484995_180557	2017-07-27 11:00:00	cassandra934a0553	10.1.1.8	90667	64.72.64.10	53	3
140484995_180558	2017-07-27 11:00:00	cassandra934a0553	10.1.1.8	90668	64.72.64.10	53	3
140484995_180559	2017-07-27 11:00:00	cassandra934a0553	10.1.1.8	90669	64.72.64.10	53	3
140484995_180560	2017-07-27 11:00:00	cassandra934a0553	10.1.1.8	90670	64.72.64.10	53	3
140484995_180561	2017-07-27 11:00:00	cassandra934a0553	10.1.1.8	90671	64.72.64.10	53	3
140484995_180562	2017-07-27 11:00:00	cassandra934a0553	10.1.1.8	90672	64.72.64.10	53	3
140484995_180563	2017-07-27 11:00:00	cassandra934a0553	10.1.1.8	90673	64.72.64.10	53	3
140484995_180564	2017-07-27 11:00:00	cassandra934a0553	10.1.1.8	90674	64.72.64.10	53	3
140484995_180565	2017-07-27 11:00:00	cassandra934a0553	10.1.1.8	90675	64.72.64.10	53	3
140484995_180566	2017-07-27 11:00:00	cassandra934a0553	10.1.1.8	90676	64.72.64.10	53	3
140484995_180567	2017-07-27 11:00:00	cassandra934a0553	10.1.1.8	90677	64.72.64.10	53	3
140484995_180568	2017-07-27 11:00:00	cassandra934a0553	10.1.1.8	90678	64.72.64.10	53	3
140484995_180569	2017-07-27 11:00:00	cassandra934a0553	10.1.1.8	90679	64.72.64.10	53	3
140484995_180570	2017-07-27 11:00:00	cassandra934a0553	10.1.1.8	90680	64.72.64.10	53	3
140484995_180571	2017-07-27 11:00:00	cassandra934a0553	10.1.1.8	90681	64.72.64.10	53	3
140484995_180572	2017-07-27 11:00:00	cassandra934a0553	10.1.1.8	90682	64.72.64.10	53	3
140484995_180573	2017-07-27 11:00:00	cassandra934a0553	10.1.1.8	90683	64.72.64.10	53	3
140484995_180574	2017-07-27 11:00:00	cassandra934a0553	10.1.1.8	90684	64.72.64.10	53	3
140484995_180575	2017-07-27 11:00:00	cassandra934a0553	10.1.1.8	90685	64.72.64.10	53	3
140484995_180576	2017-07-27 11:00:00	cassandra934a0553	10.1.1.8	90686	64.72.64.10	53	3
140484995_180577	2017-07-27 11:00:00	cassandra934a0553	10.1.1.8	90687	64.72.64.10	53	3
140484995_180578	2017-07-27 11:00:00	cassandra934a0553	10.1.1.8	90688	64.72.64.10	53	3
140484995_180579	2017-07-27 11:00:00	cassandra934a0553	10.1.1.8	90689	64.72.64.10	53	3
140484995_180580	2017-07-27 11:00:00	cassandra934a0553	10.1.1.8	90690	64.72.64.10	53	3
140484995_180581	2017-07-27 11:00:00	cassandra934a0553	10.1.1.8	90691	64.72.64.10	53	3
140484995_180582	2017-07-27 11:00:00	cassandra934a0553	10.1.1.8	90692	64.72.64.10	53	3



- Time: 9 am to 5 pm every day
- Senders: Intern's computers
- Receivers: DNS servers
- Requests: Google, Stack Overflow, various yale.edu websites

- Time: Constant and regular
- Senders: Two business computers
- Receiver: Broadcast address
- Request: Faulty printer's address
(Request denied so repeated constantly)

- Time: Between 2 am and 5 am in the morning
- Senders: Blacklisted Chinese IP addresses
- Receiver: Reservoir code repository
- Access denied

CANDID on SCinet Network

SCinet 2016 Analysis Highlights

- Network Research Exhibition (NRE)
- Metadata collected from 6 R-Scope boxes
- HPE Apollo 2000 running ENSIGN
- Data collected within specific time windows (8-24 hours)
- Data filtered by internal/external source/destination
- Binning applied by time interval, subnet, etc.
- Typical tensor ~106 non-zeros with >99% sparsity
- Typical 100 component decomposition required ~5 minutes
- Post-decomposition exploration with Splunk

Selected Fields	Actions 1	App 4	Cn 37	Dns 38	Dns_ip 18	Dns_port 6	Dropped 1	Dst 6	Eventtype 2	Filename 2	Filet 100+	Http_host 31	Http_method 3	Http_referrer 5	Http_user_agent 5	Mime_type 10	Msg 1	Note 1	P 1	Password 1	Remote_location_region 1	Rx_hosts 3	Source 8	Src 52	Src_ip 52	Src_port 100+	Status 8	Transport 2	Tunnel_greets 1	Tx_hosts 52	Uid 100+	Un 100+	Urf 100+	
	>	9/23/15 9:39:45.466 AM	54.225.146.15						>	9/23/15 9:39:44.624 AM	192.168.137.113	80	54.225.146.15																					
	>	9/23/15 9:39:43.950 AM							>	9/23/15 9:39:43.950 AM	192.168.137.113	80	54.225.146.15																					
	>	9/23/15 9:38:49.518 AM							>	9/23/15 9:38:49.518 AM	192.168.137.113	80	54.225.146.15																					
	>	9/23/15 9:38:38.614 AM							>	9/23/15 9:38:38.614 AM	192.168.137.113	5355	224.0.0.252																					
	>	9/23/15 9:38:38.514 AM							>	9/23/15 9:38:38.514 AM	192.168.137.113	5355	224.0.0.252																					
	>	9/23/15 9:38:38.514 AM							>	9/23/15 9:38:38.514 AM	192.168.137.113	5355	224.0.0.252																					
	>	9/23/15 9:38:38.500 AM							>	9/23/15 9:38:38.500 AM	192.168.137.113	67	192.168.137.1																					
	>	9/23/15 9:38:38.476 AM							>	9/23/15 9:38:29.247 AM	54.225.146.15		192.168.137.113																					
	>	9/23/15 9:38:28.940 AM							>	9/23/15 9:38:28.940 AM	192.168.137.113	80	54.225.146.15																					

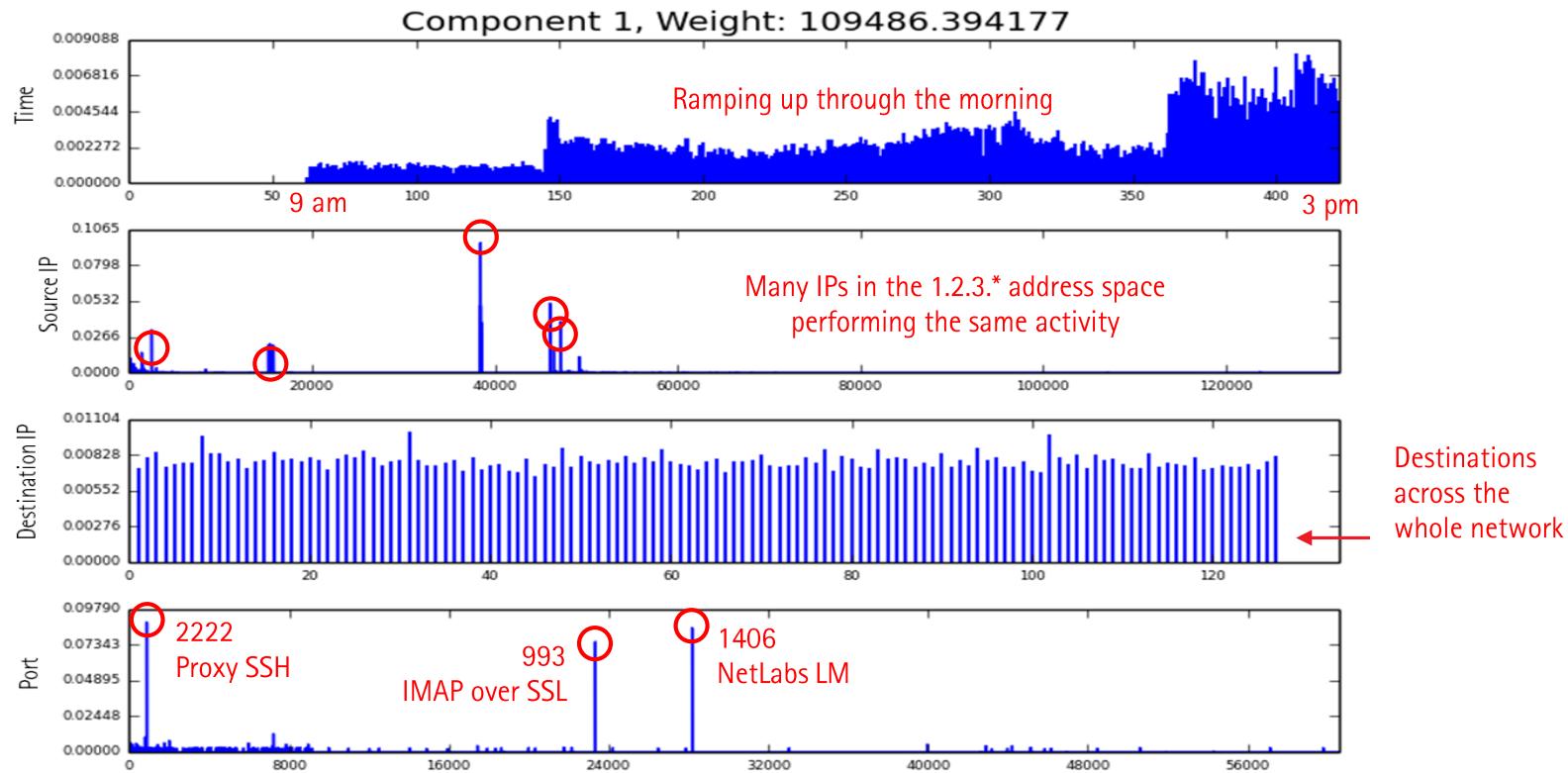
Tensors formed from selected, binned metadata pulled from filtered R-Scope (Bro) logs



HPE Apollo 2000
12 Cores, 256 GB RAM utilized per tensor

Case Study #1: External Scanners

Indicator of a reconnaissance phase of an attack

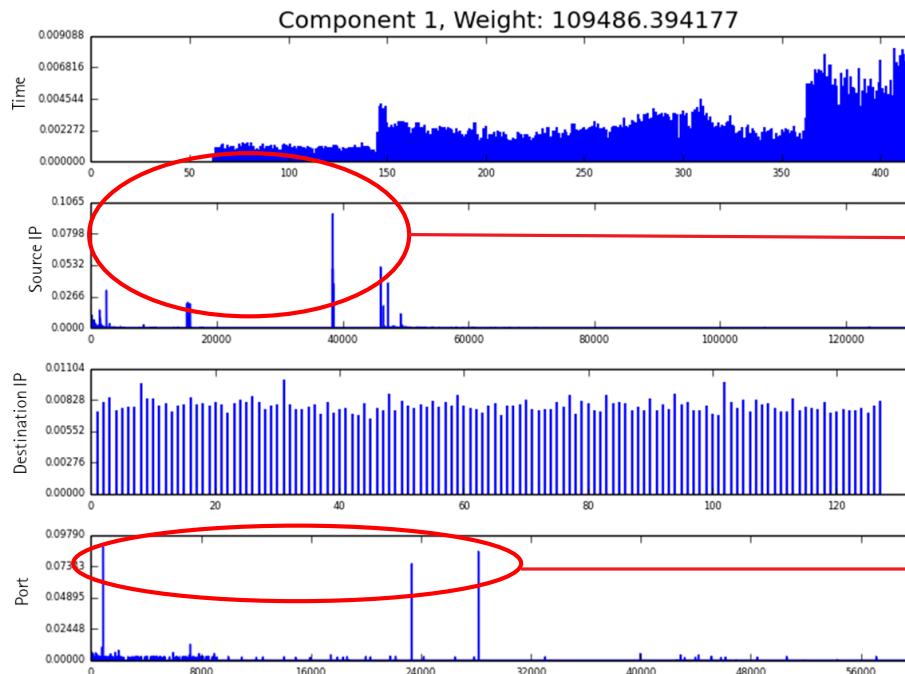


Distributed network mapping and port scanning with likelihood of hostile intent

- Coordinated attempt by multiple external actors to find hosts on SCinet with particular services

Case Study #1: External Scanners

Confirmation of hostile intent



splunk > App: Search & Reporting

Search Pivot Reports Alerts Dashboards

New Search

index="rscope" sourcetype=bro_conn id_orig_h="1.2.3.*" id_resp_p="2222"

conn_state_meaning

13 Values, 100% of events

Selected Yes No

Reports

Top values Top values by time Events with this field Rare values

scanning

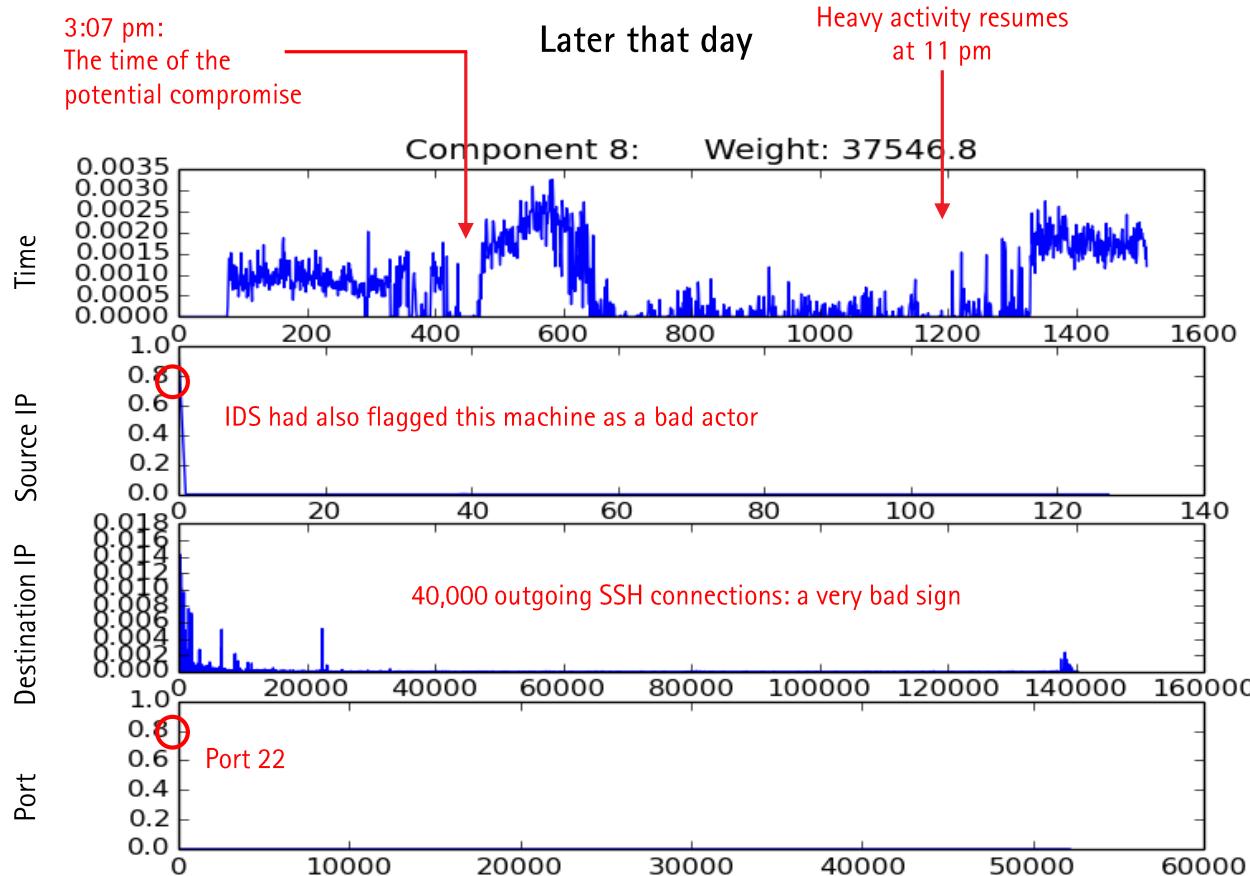
Top 10 Values	Count	%
Connection attempt seen, no reply.	19,497	87.911%
Normal establishment and termination.	1,643	7.408%
Connection established, originator aborted (sent a RST).	529	2.385%
No SYN seen, just midstream traffic (a 'partial connection' that was not later closed).	130	0.586%
Connection established, not terminated.	116	0.523%
Originator sent a SYN followed by a RST, we never saw a SYN-ACK from the responder.	90	0.406%
Responder sent a SYN ACK followed by a FIN, we never saw a SYN from the originator.	83	0.374%
Responder sent a SYN ACK followed by a RST, we never saw a SYN from the (purported) originator.	48	0.216%
Originator sent a SYN followed by a FIN, we never saw a SYN ACK from the responder (hence the connection was 'half open').	24	0.108%
Connection established and close attempt by responder seen (but no reply from originator).	7	0.032%

Further Investigation in Splunk

- Filtered search for successful connections
 - “focused” Splunk query guided by the component
- Confirmed the scanning

Case Study #1: External Scanners

Confirmation of evolution of an attack

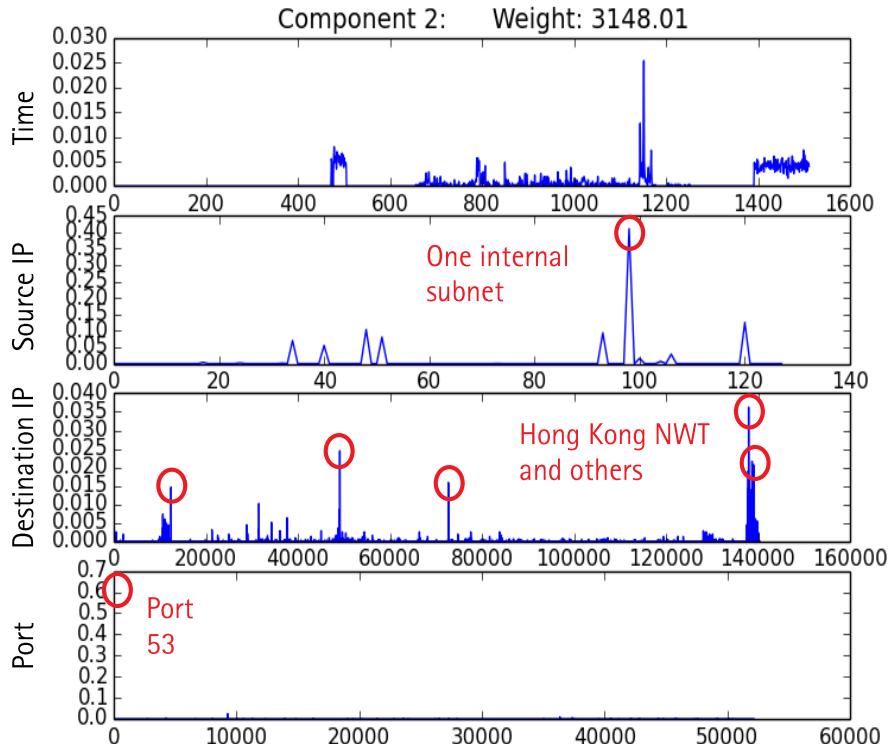


Confirmed the evolution of an attack using a later component

- Outgoing SSH connections from a compromised host

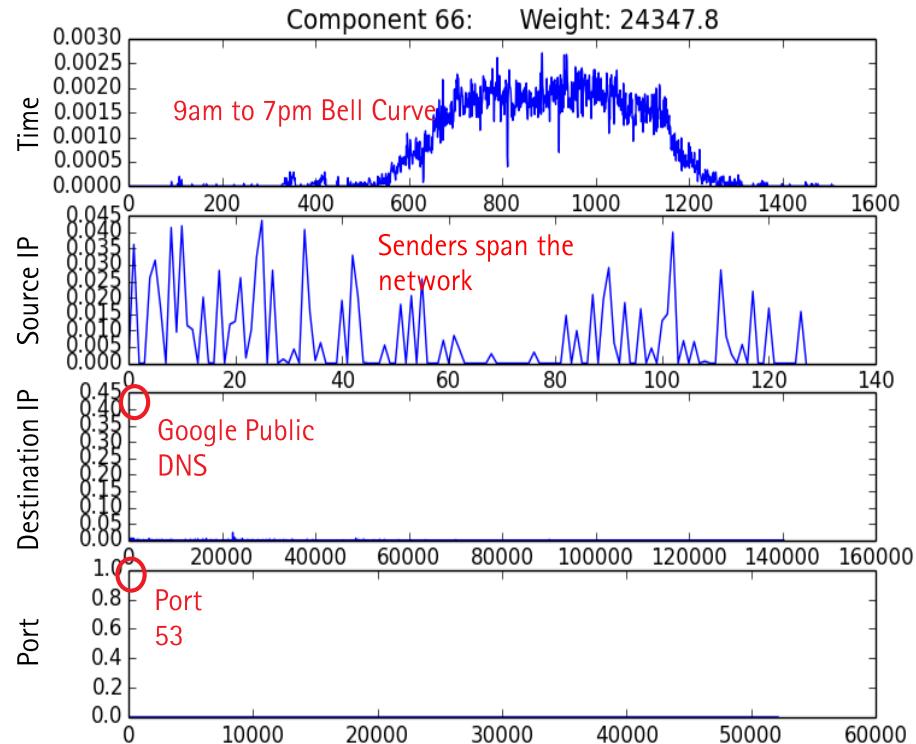
Case Study #2: Suspected Data Exfiltration

Isolating Suspicious DNS Traffic



Suspicious DNS Traffic

- Irregular spikes in time
- Suspicious destinations



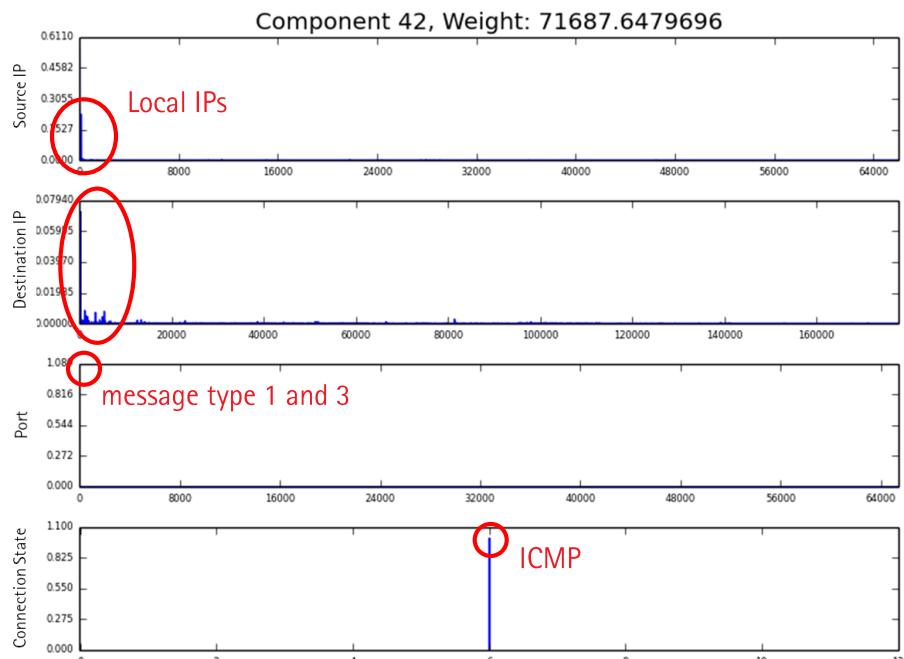
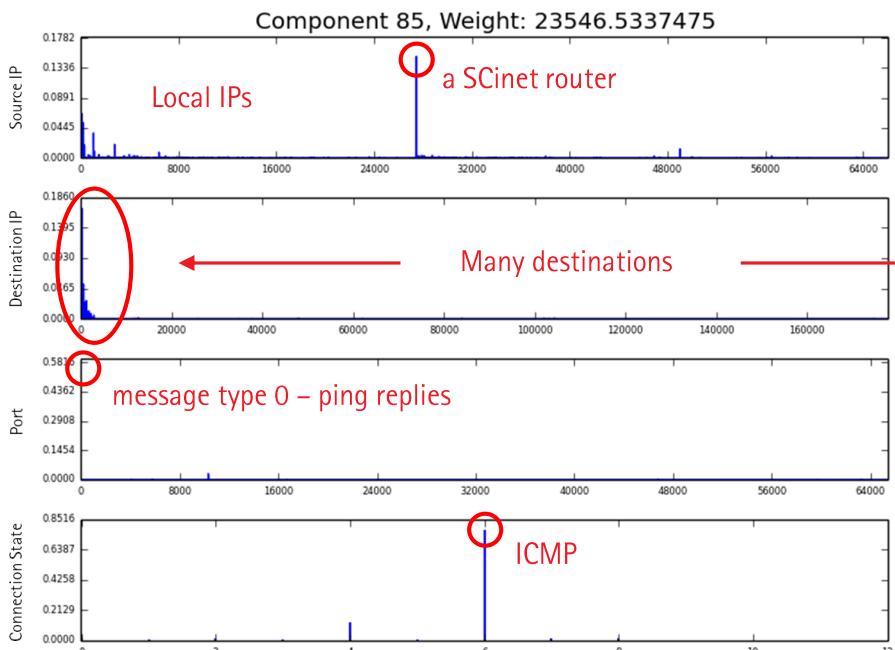
Typical DNS Traffic

- Highly regular traffic between 9 am and 7 pm – the running hours of the conference
- Valid DNS server destination

Case Study #3: ICMP Tunneling

Anomalous ICMP is difficult to distinguish through

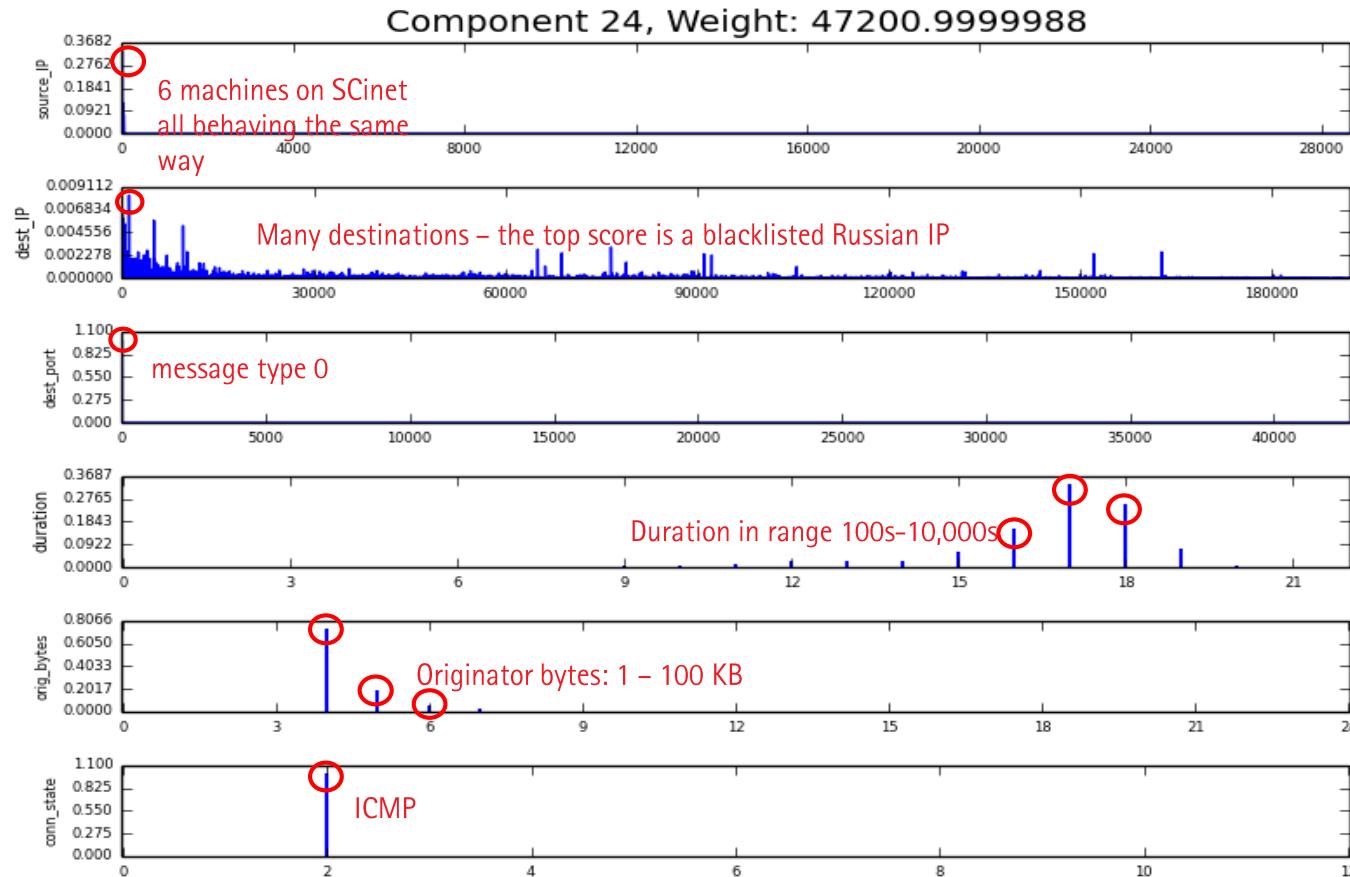
- Decompositions with fewer metadata attributes
 - IP addresses, port, connection state/time



Case Study #3: ICMP Tunneling

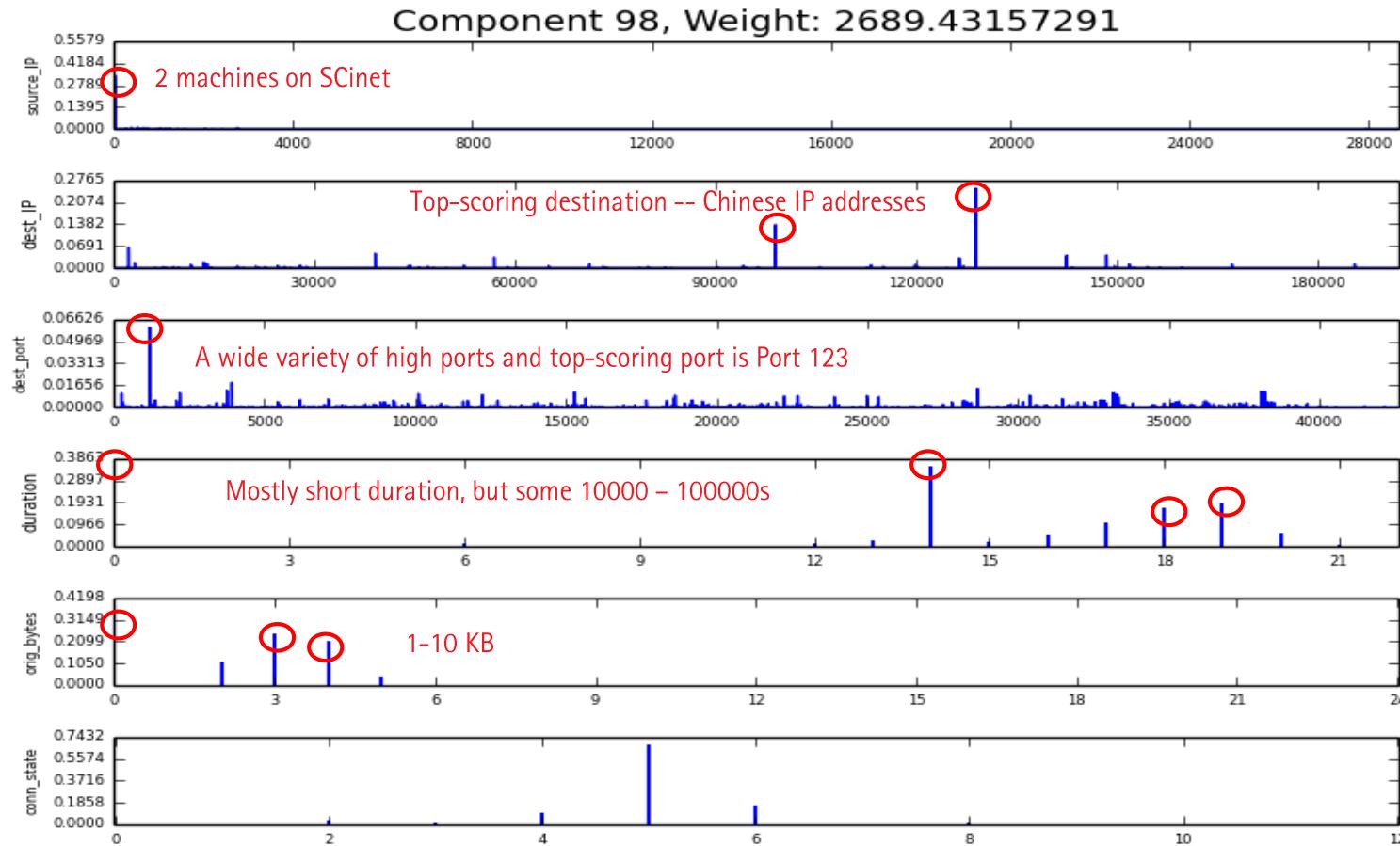
Anomalous ICMP easily distinguished through

- Decompositions with additional key metadata attributes
 - adding connection duration and number of bytes to analysis



Case Study #4: NTP Amplification Attack

Another successful use case of decompositions with more metadata attributes



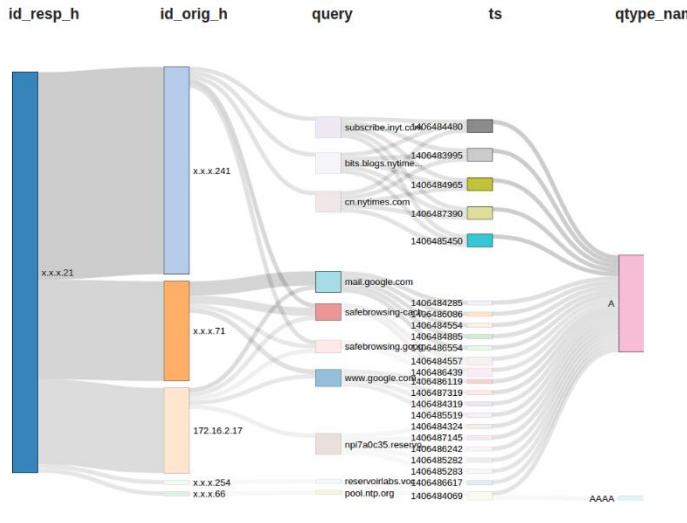
Using CANDID and ENSIGN Tensor Analysis...

We have uncovered and visualized patterns indicative of:

- Distributed port scans evolving to machine takeover
- Distributed denial of service attacks
- DNS-based data exfiltration/insider threat
- SSH password guessing (apart from scanning)
- Network policy violations
- Exploitation of application-specific port vulnerabilities
- Patterns of traffic indicative of scans for printers or IoT devices
- Broken or misconfigured network services
- Selective, persistent use of cryptographic methods in point-to-point communication

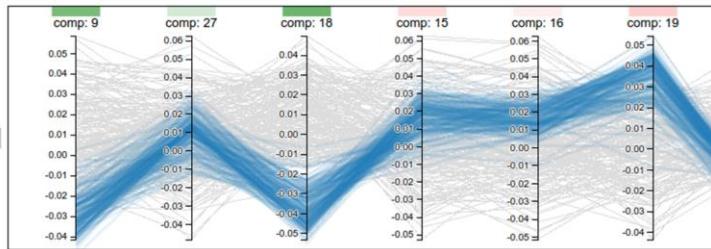
```
01/15/2015-18:15:58.890499  [**] [1:2020159:6] ET CURRENT_EVENTS Upatre Redirector Jan 9 2015  
[*] [Classification: A Network Trojan was detected] [Priority: 1] {TCP} 172.16.120.154:49380 ->  
86.35.15.212:80  
01/15/2015-18:17:21.889077  [**] [1:2003492:19] ET MALWARE Suspicious Mozilla User-Agent - Like  
ly Fake (Mozilla/4.0) [*] [Classification: A Network Trojan was detected] [Priority: 1] {TCP}  
172.16.120.154:49407 -> 202.153.35.133:20110
```

More Features in the Pipeline

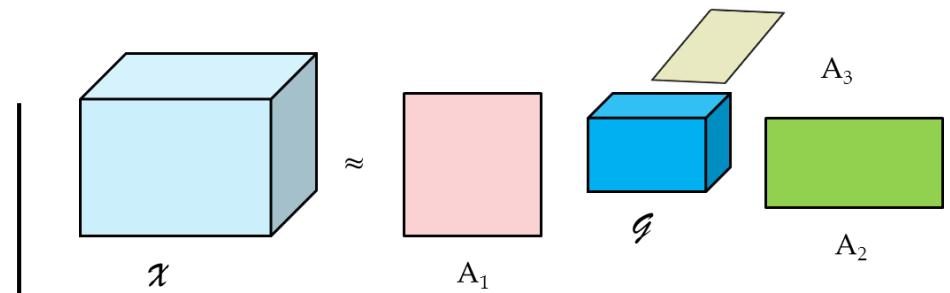


Correlation

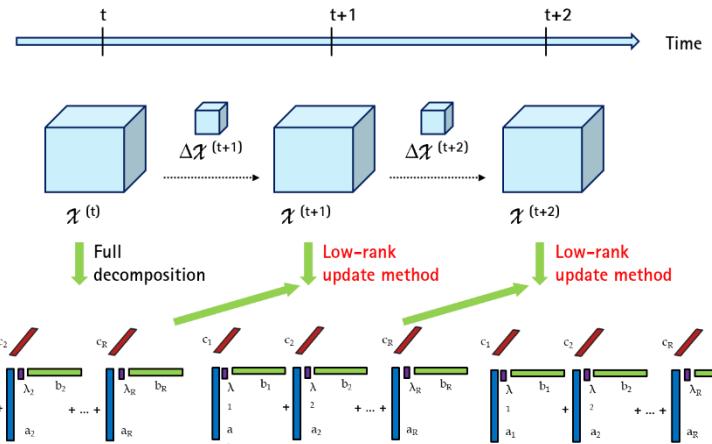
cluster 0: size 168 ▾ Show Related



Specialized visualization and
ML-driven post-processing methods



Alternate and advanced methods for tensor decompositions



Support for streaming updates

Conclusion

Contact **Reservoir** Labs

- <https://www.reservoir.com>

Contact the Speaker

- baskaran@reservoir.com

Meet us at SCinet NRE Demo 2017

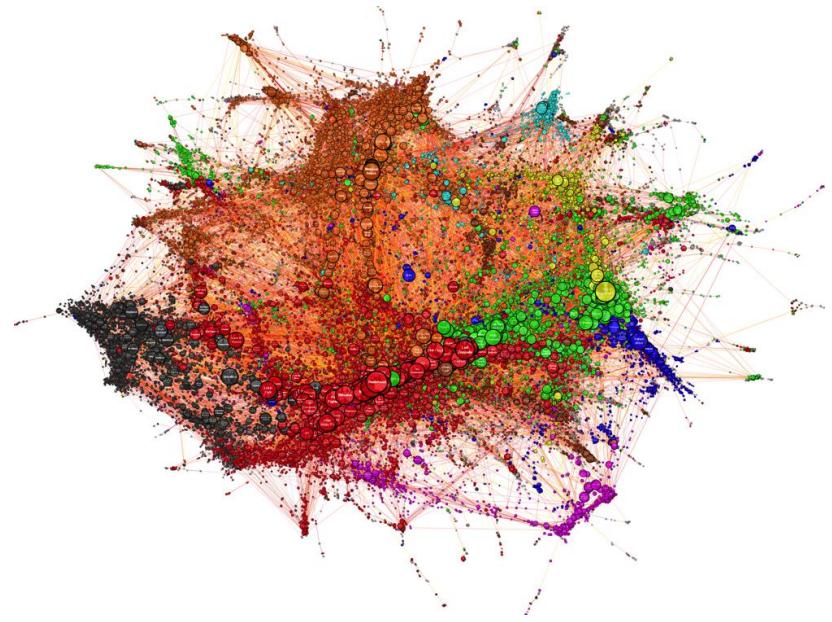
Other Recent Papers

- *Cyber Security Through Multidimensional Data Decompositions*

D. Bruns-Smith, M. Baskaran, T. Henretty, J. Ezick,
R. Lethin, in CYBERSEC, Apr 2016

- *Memory-efficient Parallel Tensor Decompositions*

M. Baskaran, T. Henretty, D. Bruns-Smith, M. H.
Langston, J. Ezick, R. Lethin, in IEEE HPEC, Sep 2017
(Best Paper Award)



Tensor decompositions provide a fast, scalable linear algebra based solution to finding patterns in linked metadata