Dynamic Factor Graphs for Attack Preemption

Phuong Cao, Zbigniew Kalbarczyk, Ravishankar Iyer

ILLINOIS
NCSA | National Center for Supercomputing Applications

U.S. National Science Foundation
CICI, CC*, FMiTF, RoRS, SaTC

Secure-HPC Workshop
Supercomputing' 25
Nov 16th, St. Louis, MO

Security Challenges in Cloud vs. HPC

Cloud

HPC

Low trust, built-in isolation
- Virtual Private Cloud
- Security Groups
- Encrypted FS in transit/at rest

Least privilege by default
- Resource isolation

Target
- Service disruption
- Account takeover

Distributed monitoring
- Host-based agents

High Trust
- Federated authentication
- Shared file systems, GPU drivers
- Message passing

High Privilege
- Custom code modification is norm
- More sensitive and proprietary data

Target
- Data exfiltration and IP theft
- Resource hijacking

Centralized monitoring
- Bastion host, optical tap

Photos: Adobe Stock Photo, NCSA Radiant

Security Challenges in Cloud vs. HPC

Cloud

HPC

Low trust, built-in isolation
- Virtual Private Cloud
- Security Groups
- Encrypted FS in transit/at rest

High Trust
- Federated authentication
- Shared file systems, GPU drivers
- Message passing

Least privilege by default
- Resource isolation

High Privilege
- Custom code modification is norm
- More sensitive and proprietary data

Target
- Service disruption
- Account takeover

Target
- Data exfiltration and IP theft
- Resource hijacking

Distributed monitoring
- Host-based agents

Centralized monitoring
- Bastion host, optical tap

Photos: Adobe Stock Photo, NCSA Radiant

Security Challenges in Cloud vs. HPC

Cloud

HPC

Low trust, built-in isolation
- Virtual Private Cloud
- Security Groups
- Encrypted FS in transit/at rest

Least privilege by default
- Resource isolation

High Trust
- Federated authentication
- Shared file systems, GPU drivers
- Message passing

High Privilege
- Custom code modification is norm
- More sensitive and proprietary data

Target
- Service disruption
- Account takeover

Target
- Data exfiltration and IP theft
- Resource hijacking

Distributed monitoring
- Host-based agents

Centralized monitoring
- Bastion host, optical tap

Photos: Adobe Stock Photo, NCSA Radiant

Security Challenges in Cloud vs. HPC

Cloud

Low trust, built-in isolation
- Virtual Private Cloud
- Security Groups
- Encrypted FS in transit/at rest

Least privilege by default
- Resource isolation

Target
- Service disruption
- Account takeover

Distributed monitoring
- Host-based agents

HPC

High Trust
- Federated authentication
- Shared file systems, GPU drivers
- Message passing

High Privilege
- Custom code modification is norm
- More sensitive and proprietary data

Target
- Data exfiltration and IP theft
- Resource hijacking

Centralized monitoring
- Bastion host, optical tap

Photos: Adobe Stock Photo, NCSA Radiant

Summary of Dynamic Factor Graphs for Attack Preemption (1)

Challenges

- Overwhelming noise and alert volumes
- Relying on late-state critical alerts
- Necessity of alert sequence detection

State of the Art

- Atomic alert-based system
- Black Hole Routing
- Deep generative models

Summary of Dynamic Factor Graphs for Attack Preemption (1)

Challenges

- Overwhelming noise and alert volumes
- Relying on late-state critical alerts
- Necessity of alert sequence detection

State of the Art

- Atomic alert-based system
- Black Hole Routing
- Deep generative models

Our Contributions

- **Characterization** of alert sequences in real-world, national scale HPC providing GPU resources
- **Insights:**

(1) Critical alerts arriving late

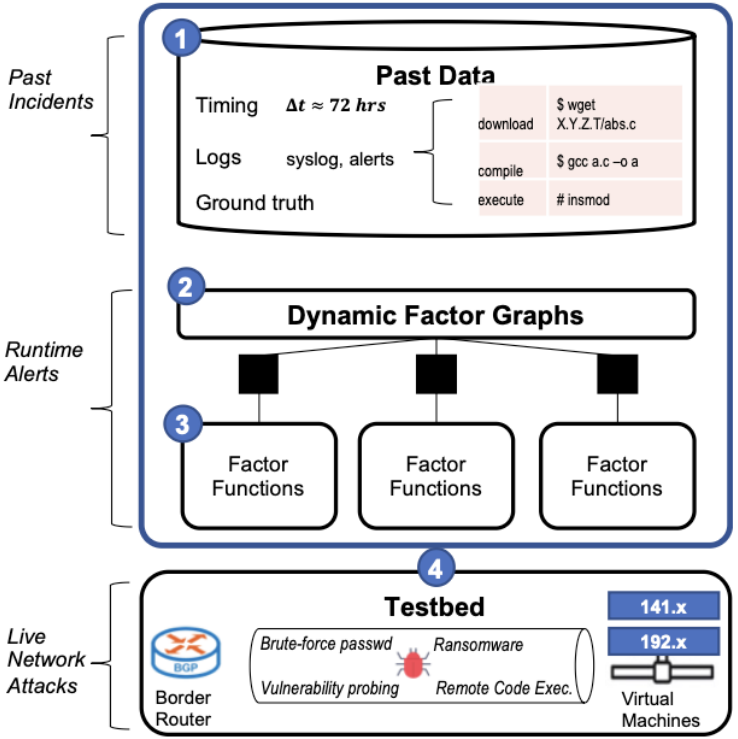
(2) Alert sequences length from two to four for detection

(3) Similarity of attacks (95% show 33% similar sequences)

(4) Timings of recurrent alerts matter.

- **Dynamic Factor Graphs:**

Uncertainty, conditional probabilities, and evolution of alerts



Summary of Dynamic Factor Graphs for Attack Preemption (2)

Challenges

- Overwhelming noise and alert volumes
- Relying on late-state critical alerts
- Necessity of alert sequence detection

State of the Art

- Atomic alert-based system
- Black Hole Routing
- Deep generative models

Our Contributions

- **Characterization** of alert sequences in real-world, national scale HPC providing GPU resources
- **Insights:**

(1) Critical alerts arriving late

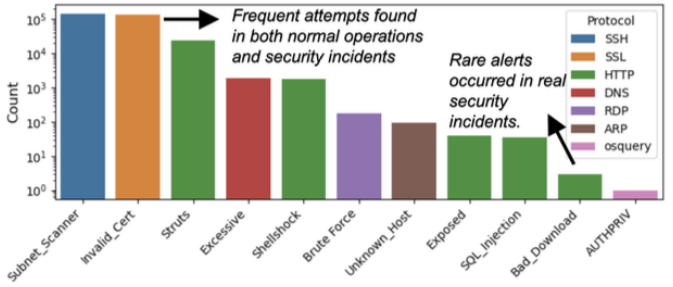
(2) Alert sequences length from two to four for detection

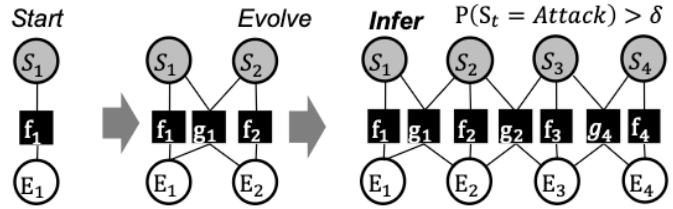
(3) Similarity of attacks (95% show 33% similar sequences)

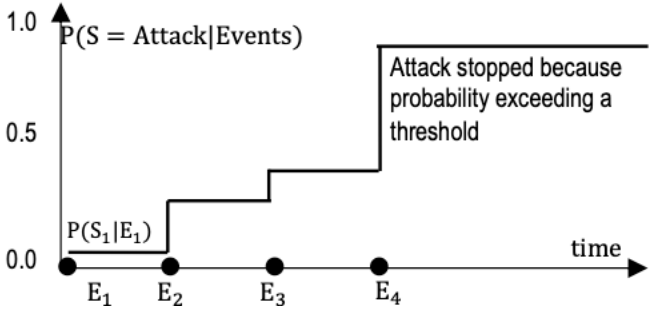
(4) Timings of recurrent alerts matter.

- **Dynamic Factor Graphs:**

Uncertainty, conditional probabilities, and evolution of alerts







Summary of Dynamic Factor Graphs for Attack Preemption (3)

Challenges

- Overwhelming noise and alert volumes
- Relying on late-state critical alerts
- Necessity of alert sequence detection

State of the Art

- Atomic alert-based system
- Black Hole Routing
- Deep generative models

Our Contributions

- **Characterization** of alert sequences in real-world, national scale HPC providing GPU resources
- **Insights:**

(1) Critical alerts arriving late

(2) Alert sequences length from two to four for detection

(3) Similarity of attacks (95% show 33% similar sequences)

(4) Timings of recurrent alerts matter.

- **Dynamic Factor Graphs:**

Uncertainty, conditional probabilities, and evolution of alerts

Future Work

- **Providing data on emerging vulnerabilities**

0-day in drivers of accelerators (GPU, QPU)

Provenance of AI models (data, code)

Quantum-resistant cryptography pef. overhead

MCP honeypot

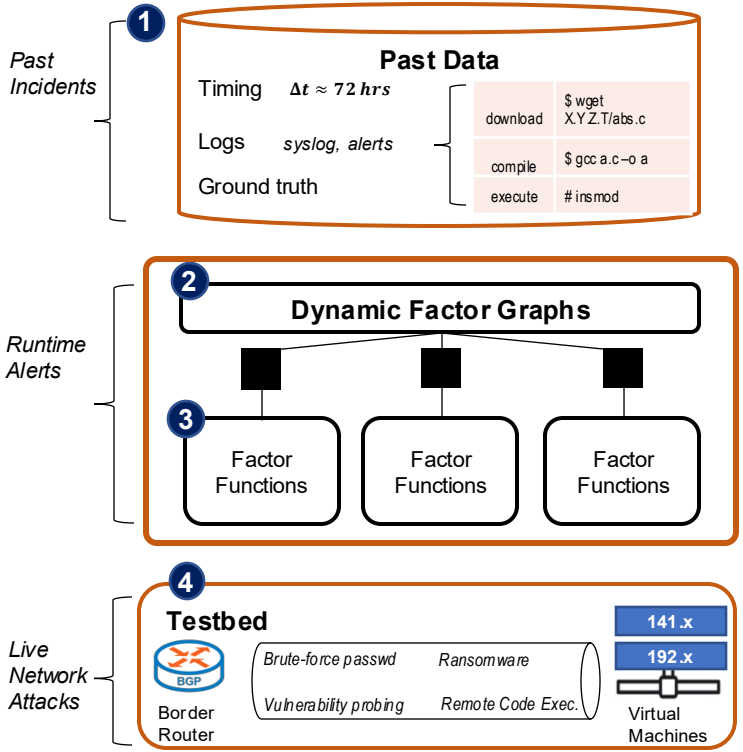
- **Focus on attack recovery**

Automated integration with Black Hole Router

Minimize workload disruptions via checkpointing

- **New research security policies on AI/HPC**

Part I. Data-Driven Insights

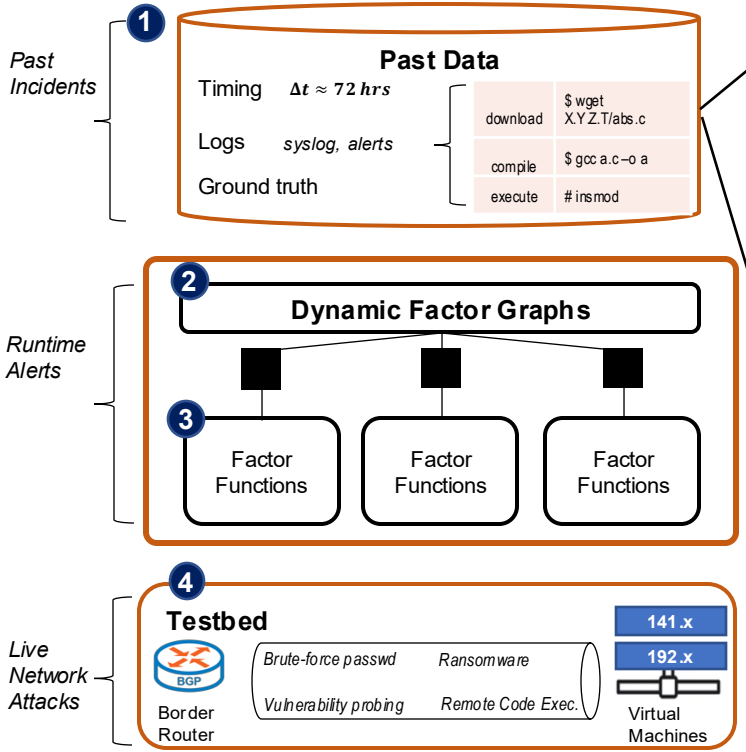


Data driven insights on alerts preceding attacks

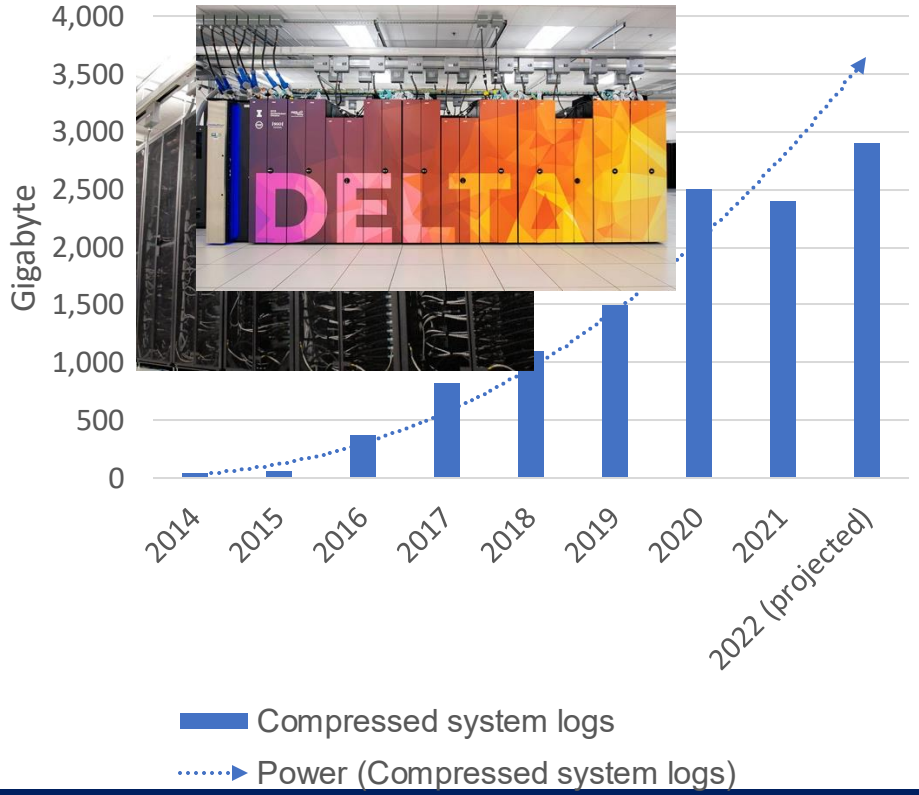
Dynamic Factor Graphs of factor functions

Results preemption

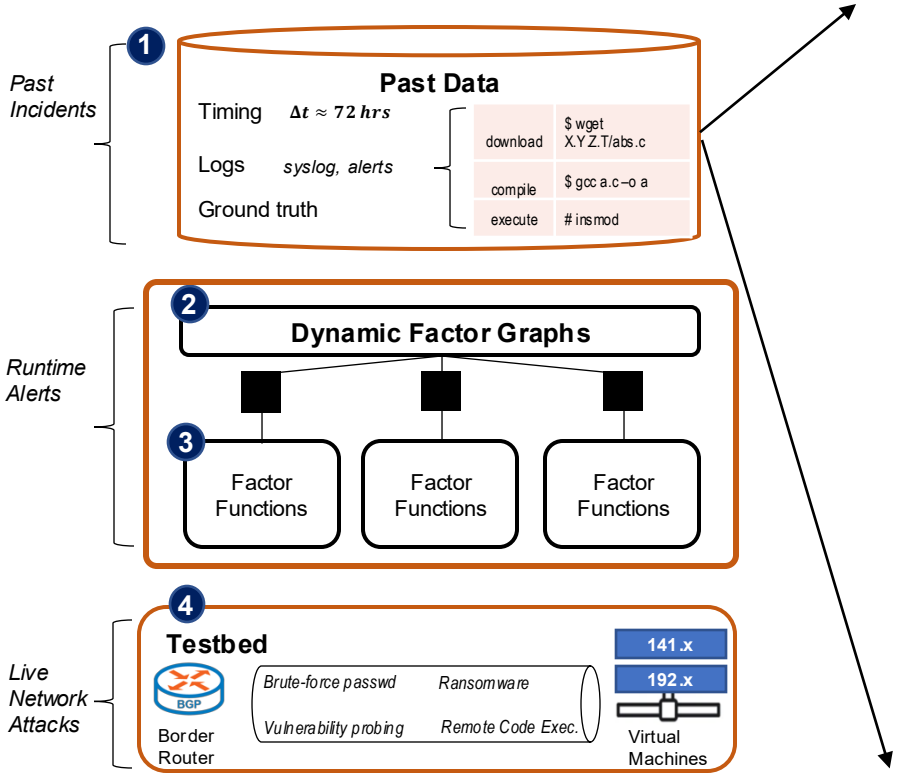
Motivation for Data Curation



Longitudinal, growing resiliency and security dataset



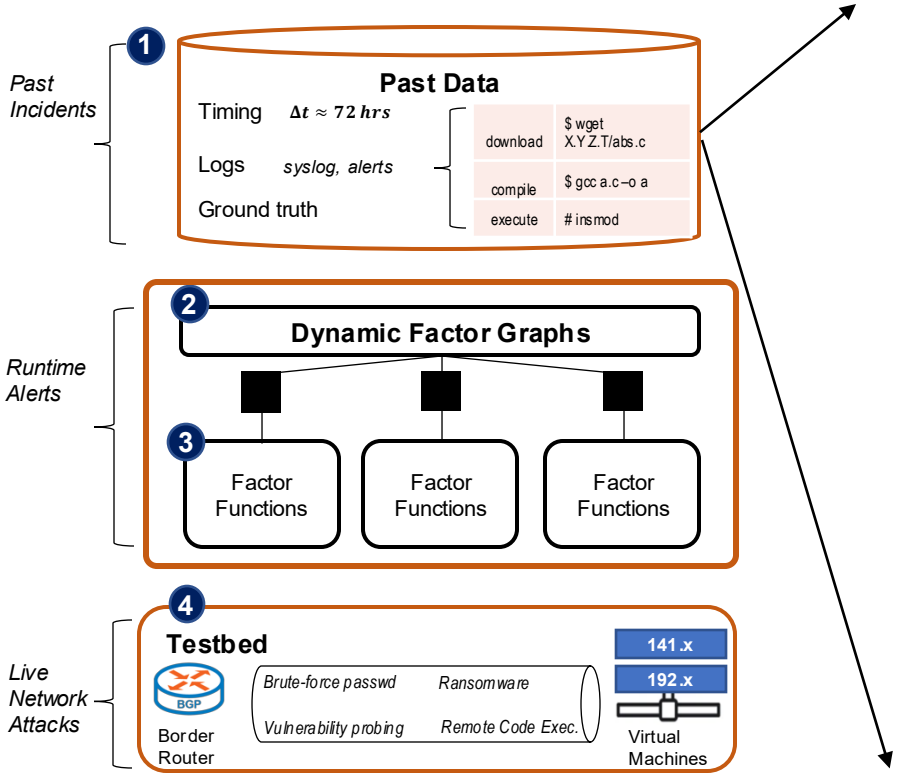
Data Collection Statistics



Data type	Summary statistics
Number of hosts	5,000+ (clusters, workstations, laptops)
Number of active users	6,000+
Network	Class B (/16) up to 65,535 IP addresses
Network link	4.5 Tbps
Monitoring data	Zeek (4.5 GB daily) Central syslog (1.5 GB daily) Persistent logs (20 TB total)
OS types	Linux, Windows, macOS

Table 1. Summary of security log data.

Data Collection System



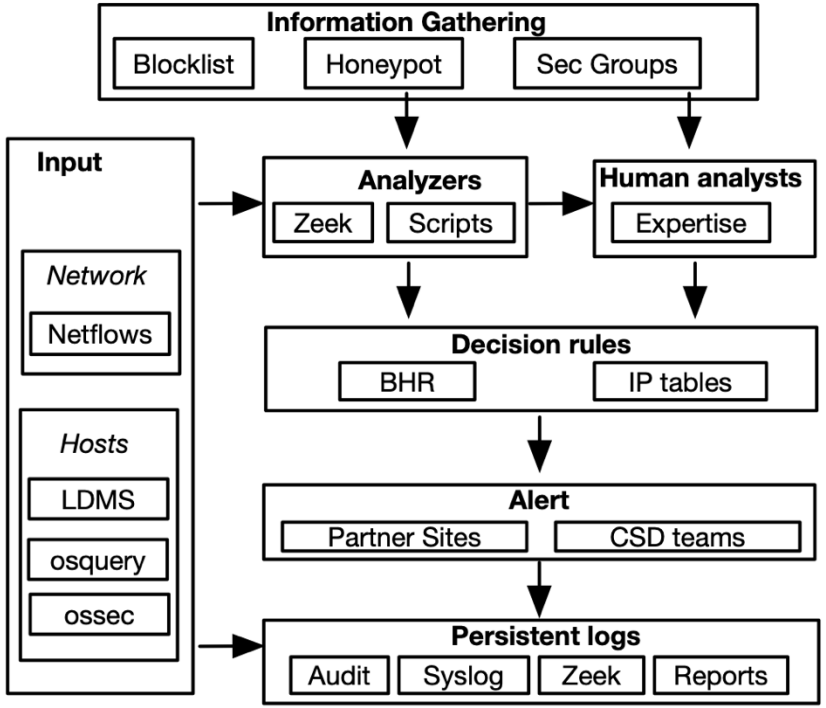
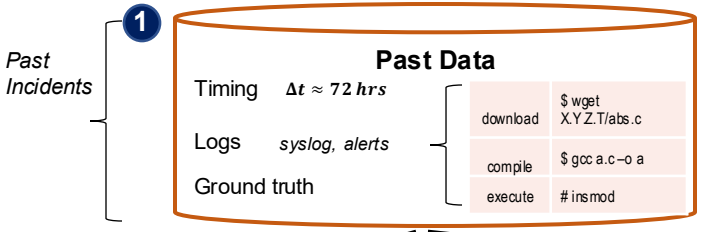
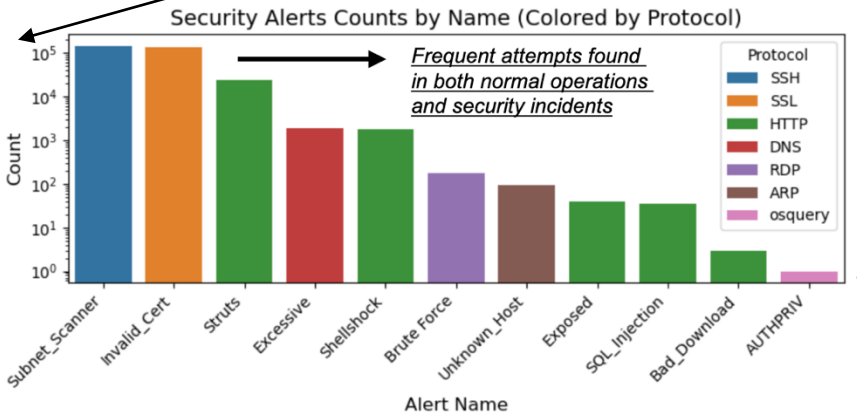


Figure 2. Overview of security monitoring tools.

Alert Distribution: Common vs. Rare Alert



Rare Alerts occurred in real security incident is a strong indicator of an attack -> Quantifying uncertainty of critical alert is needed.

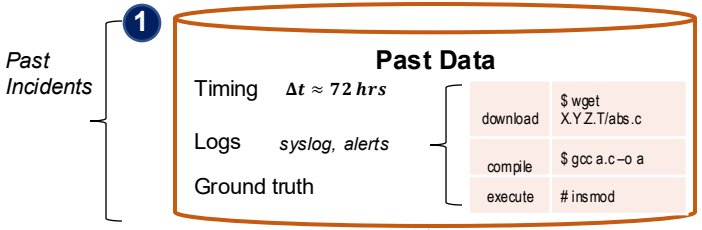


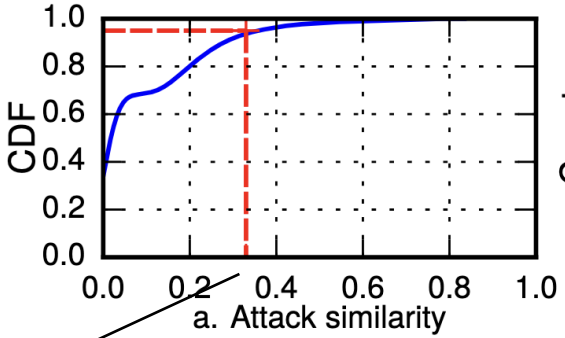
Alert Description

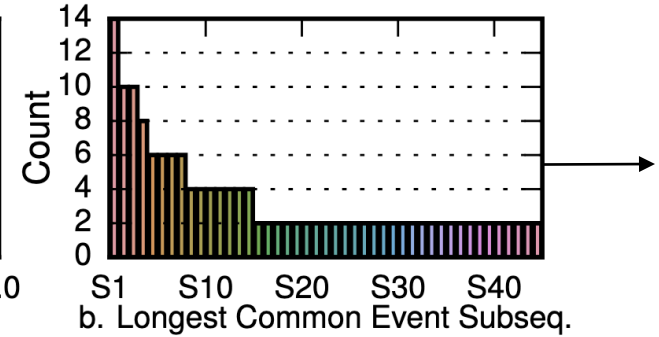
Count	Service	Name	Description/Example
145.9K	SSH	Subnet_Scanner	Scan for SSH hosts
133.8K	SSL	Invalid_Cert	Invalid server certificate
23.4K	HTTP	Struts	Exploit Apache Struts
1.9K	DNS	Excessive	Large outgoing DNS requests
1.8K	HTTP	Shellshock	Attempt to exploit Bash
175	RDP	Brute Force	Remote desktop login
93	ARP	Unknown_Host	New host on internal network
39	HTTP	Exposed	An internal server is exposed
36	HTTP	SQL_Injection	Inject SQL commands
3	HTTP	Bad_Download	A known malicious file
1	osquery	AUTHPRIV	Root privilege escalation

Rare alerts occurred in real security incidents.

Alert Similarities among Attacks



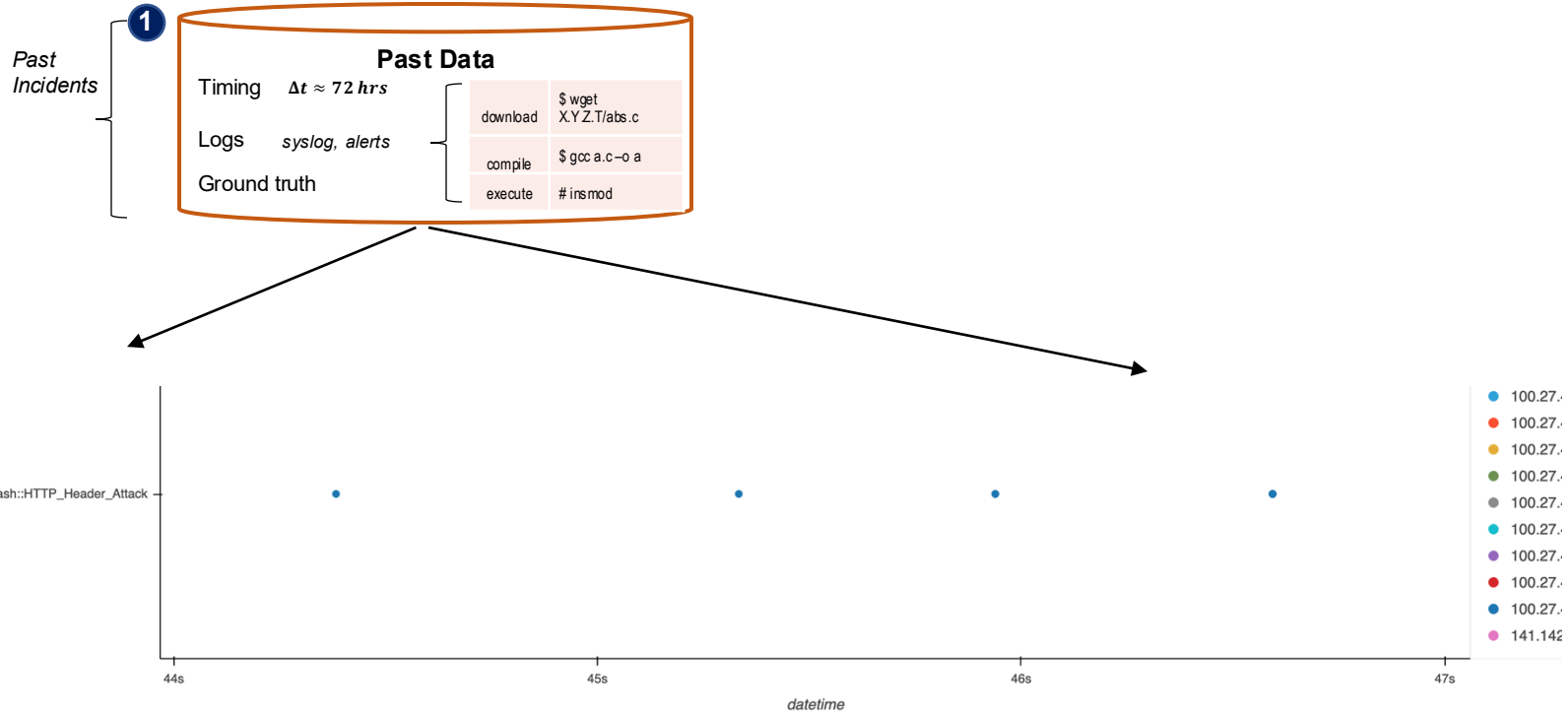




Pairs of security incidents tend to exhibit alert sequences of 2 to 5 events in length

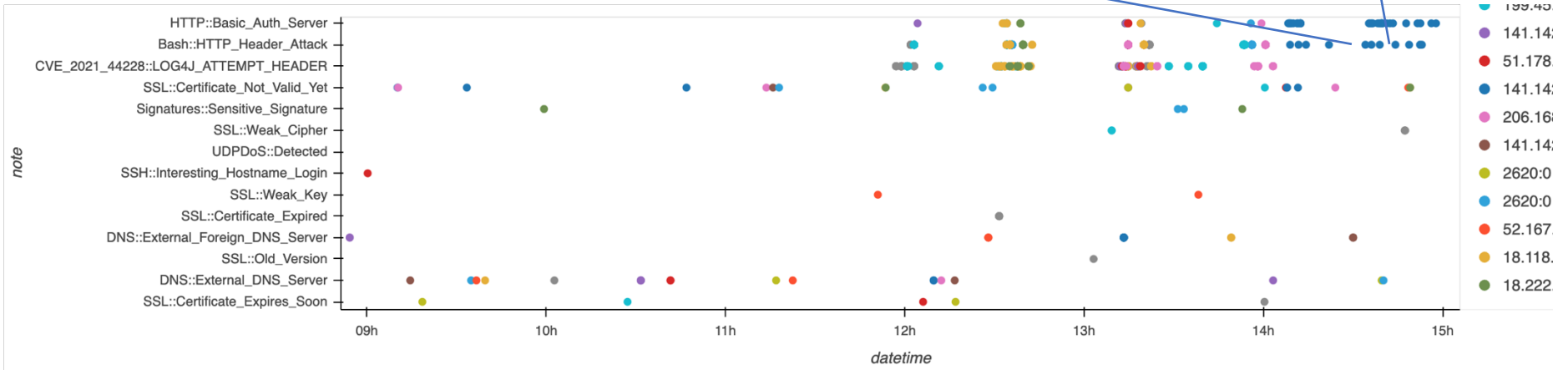
At 95% CI, any two security incidents share 33% of similar alert type.

Alert Interval Between Attack Attempts

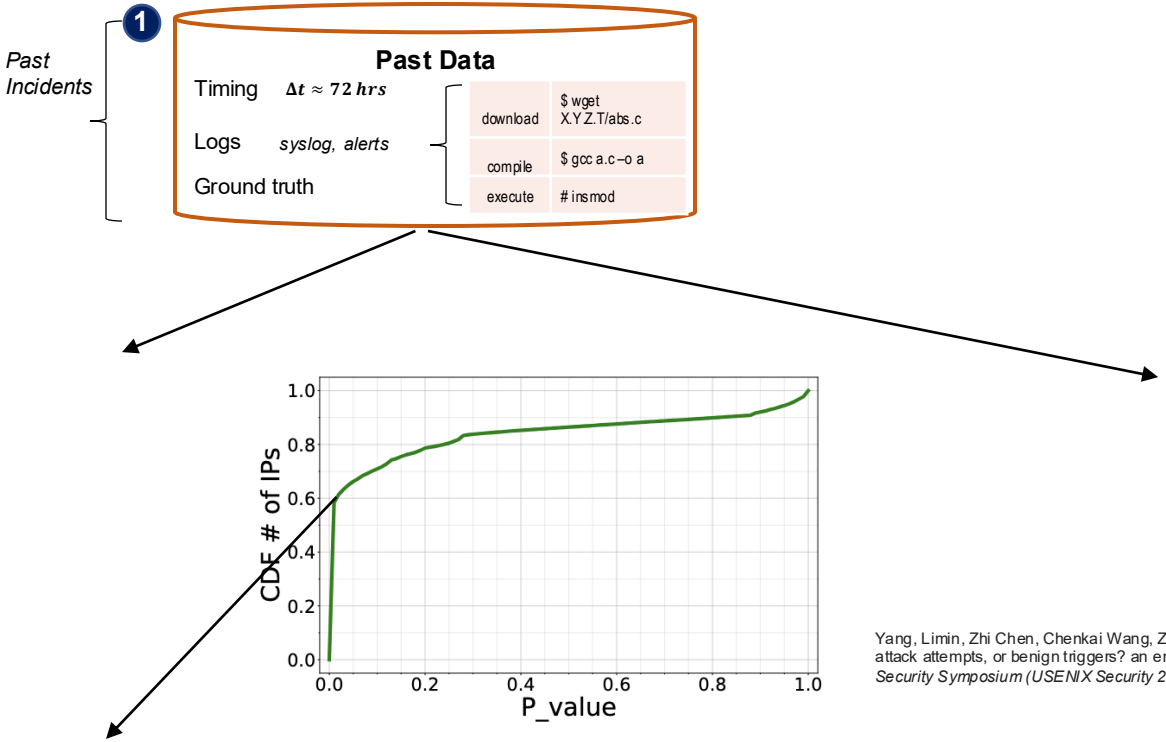


Timing and interval of alerts indicate the degree of automation (human or bot)

Repeated Alerts vs. Irregular alerts



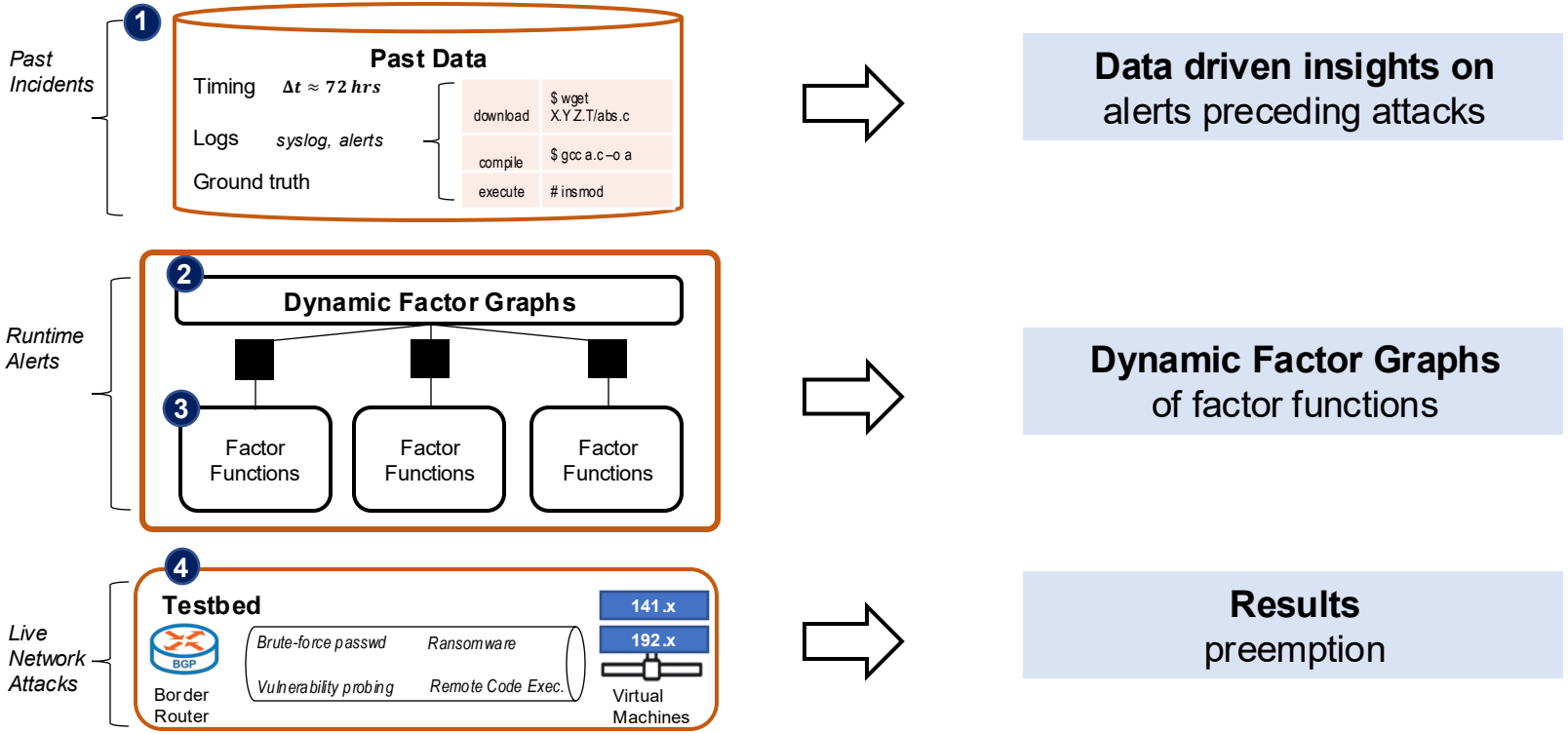
Statistical test for precisely repeated alerts



Yang, Limin, Zhi Chen, Chenkai Wang, Zhenning Zhang, Sushruth Booma, Phuong Cao, Constantin Adam et al. "True attacks, attack attempts, or benign triggers? an empirical measurement of network alerts in a security operations center." In *33rd USENIX Security Symposium (USENIX Security 24)*, pp. 1525-1542. 2024.

Testing for precise interval of alerts help filter bots vs. human-based attacks

Part II. Dynamic Factor Graphs

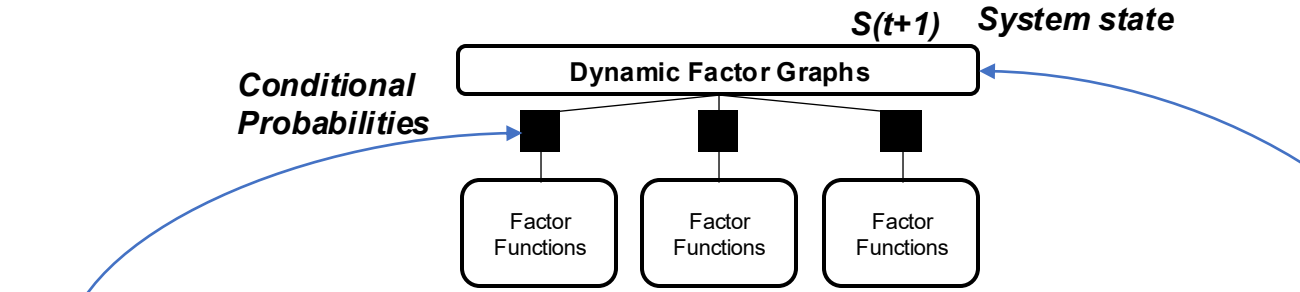


Data driven insights on alerts preceding attacks

Dynamic Factor Graphs of factor functions

Results preemption

Why Dynamic Factor Graphs?



Uncertainty

Attack attempt *vs. True Attacks*
Critical alert *vs. Regular Alert*
Human *vs. Bot*

Dynamism

Evolving attacks
Past information taken into account

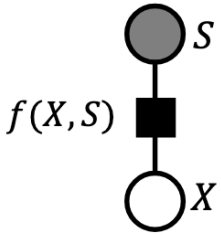
Prediction

Inferring system state and user's intent.

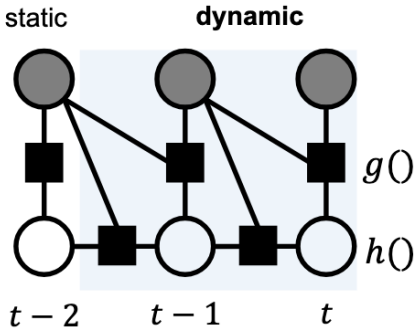
Why Dynamic Factor Graphs?

Common Features	**Bayesian Networks**	**Markov Random Fields**	**Factor Graphs**
Typical Applications	Medical diagnosis, causal modeling	Image processing, natural language processing	Coding theory, error correction, security
Graph Structure	Directed Acyclic Graph (DAG)	Undirected graphs	Bipartite (two types of nodes)
Nodes	Random variables	Random variables	Variables and factors
Edges	Directed $X \rightarrow Y$	Undirected X — Y	Undirected, e.g., $f(X,Y)$ as X — f — Y
Representation	Conditional probabilities	Potential functions	Factors (any multivariate function)
Inference	Variable elimination, junction tree	Mean-field, Gibbs sampling	Belief propagation, sum-product, max-product
Unique Features	**Bayesian Networks**	**Markov Random Fields**	**Factor Graphs**
Complex dependencies	Conditional dependency between immediate variables such as a parent and a child.	First order Markov property depending only on immediate previous state (e.g., memoryless)	✓ *Model higher-order and multiple complex dependencies, e.g., alert sequences in attacks.*
Explicit structure	A BN typically shows conditional probabilities $P(Y\|X)$ but does not explicitly specify prior $P(X)$ or $P(Y)$ in a BN.	A MRF shows potential functions of a clique of variables, e.g., $P(X,Y,\ldots,Z,T)$, but does not specify pairwise variable relationships.	✓ *Explicitly specify a prior probability $P(X)$, a conditional probability $P(Y\|X)$ or a potential function $P(X,Y)$ using factor functions.*
Unified Representation	Only directed dependencies	Undirected dependencies	✓ *Unify representation of directed (BN) and undirected (MRF) using bipartite graph.*
Efficient inference	Exact inference on small BN models	Approximate inference using Gibbs sampling	✓ *Fast inference using Belief Propagation or approximation using Gibbs sampling.*

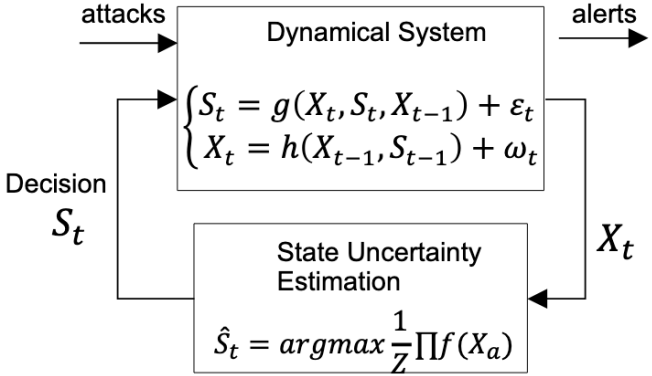
Dynamic Factor Graphs



a) A static Factor Graph of a single state variable



b) A dynamic Factor Graph (DFG) evolved from a static FG



c) Dynamic Equations of inference in a DFG

Factor Functions

$$\underbrace{P(S_{1..n}, E_{1..n})}_{\text{joint distribution}} = \frac{1}{Z}\prod_{i=1}^{n}\left[\underbrace{f(S_i, E_i)}_{\text{severity}} \times \underbrace{g(S_i, E_1, .., E_t)}_{\text{similarity}} \times \underbrace{h(S_i, E_t)}_{\text{sophistication}}\right]$$

$$S = \{benign, suspicious, \boldsymbol{preemptive}, malicious\}$$

Safe state Preemptive **Unsafe state**

Factor Functions

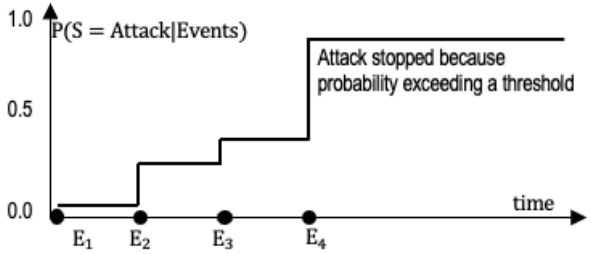
	Factor Function	Relation to attack state	Definition	Notations
Severity	$f(S_t, \Phi(E_t))$	Measures the severity indicated by a single alert to assess attack state	$e^{-\lambda_{i,f}\delta(S_t,\Phi(E_t))^2}$	$\Phi : \mathcal{E} \mapsto \mathcal{S}$ maps an alert to an attack state based on the occurrence of the alert in the past (real) attacks and in normal (benign) operations λ_f: rate of severity
Similarity	$g(S_t, \Phi(S_t, E_i, E_j, \ldots, E_t))$	Measures the similarity of an alert sequence for resemblance to past attacks	$e^{-\lambda_{i,f}\delta(S_t,\Phi(E_i,E_j,\ldots,E_t))^2}$	$\Phi_n : \mathcal{E}^n \mapsto \mathcal{S}$ maps an alert sequence ($n \geqq 2$) to an attack state based on presence of the sequence in the past (real) attacks and in normal (benign) operations. λ_g: rate of similarity
Sophistication	$h(S_t, \Phi(E_t))$	Measures the attack complexity by analyzing repetitiveness, distinguishing predictable bot patterns from sophisticated human tactics.	$e^{-\lambda_h\delta(S_t,\Phi(E_t))^2}$	$\Phi : \mathcal{E} \mapsto \mathcal{S}$ maps an alert E to an attack state based on autocorrelation to detect periodic patterns and randomness in alert triggers. λ_h: rate of sophistication

The likelihood of alerts in relation to attacks

- **Identify sequences of alerts***

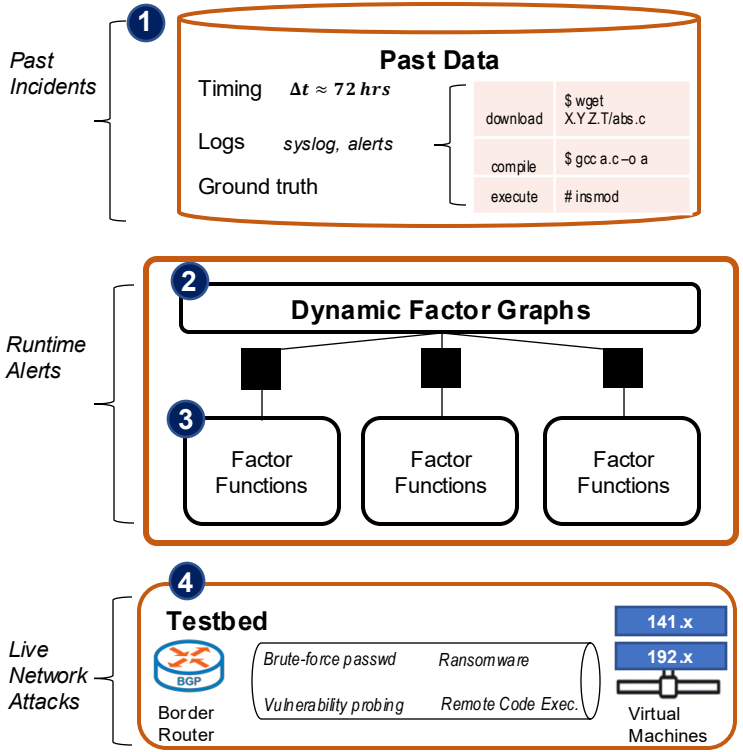
- **Hypothesis Testing:**
  - **Null Hypothesis (H_0): The alert sequence is not indicative of an attack.**
  - **Alternative Hypothesis (H_1): The alert sequence is indicative of an attack.**

- **Test Statistic:**
  - **Prior probability: Probability that an individual alert is an indicator of an attack**
  - **Posterior Probability: The posterior probability of an attack given progressively more alerts**
    - **P(Attack | A1) P(Attack | A2) P(Attack | A3)**



*using Longest common subsequence (LCS) algorithm

Part III. Results



Data driven insights on alerts preceding attacks

Dynamic Factor Graphs of factor functions

Results preemption

Alert Patterns

Alert 1	Alert 2	Alert 3
ALERT_SENSITIVE_HTTP_URI	ALERT_SENSITIVE_HTTP_URI	
ALERT_INTERNAL_ADDRESS_SCAN	ALERT_INTERNAL_ADDRESS_SCAN	
ALERT_MULTIPLE_LOGIN	ALERT_SENSITIVE_HTTP_URI	
ALERT_MULTIPLE_LOGIN	ALERT_WEAK_PASSWORD_LOGIN	
ALERT_ANOMALOUS_HOST	ALERT_SENSITIVE_HTTP_URI	ALERT_INSTALL_BOT
ALERT_SENSITIVE_HTTP_URI	compile	
ALERT_ANOMALOUS_HOST	ALERT_WEAK_PASSWORD_LOGIN	
ALERT_HIGH_NETWORKFLOWS	ALERT_NEW_SENSITIVE_CONNECTION	
ALERT_SENSITIVE_HTTP_URI	ALERT_MALWARE_HASH_REGISTRY_MATCH	
ALERT_MALWARE_HASH_REGISTRY_MATCH	ALERT_SENSITIVE_FTP_URI	
ALERT_MULTIPLE_LOGIN	ALERT_INTERNAL_ADDRESS_SCAN	
ALERT_MULTIPLE_LOGIN	ALERT_MALWARE_HASH_REGISTRY_MATCH	ALERT_SENSITIVE_FTP_URI
ALERT_SSH_BRUTEFORCE	ALERT_MULTIPLE_LOGIN	ALERT_WEAK_PASSWORD_LOGIN
ALERT_MULTIPLE_LOGIN	read_host_configuration	
ALERT_HIGH_NETWORKFLOWS	ALERT_NEW_SENSITIVE_CONNECTION	
ALERT_NEW_IRC_CONNECTION	ALERT_SENSITIVE_HTTP_URI	
ALERT_ANOMALOUS_HOST	ALERT_SENSITIVE_HTTP_URI	ALERT_INSTALL_BOT
ALERT_WATCHED_COUNTRY_LOGIN	ALERT_SENSITIVE_HTTP_URI	
ALERT_SENSITIVE_HTTP_URI	ALERT_MALWARE_HASH_REGISTRY_MATCH	
ALERT_SENSITIVE_HTTP_URI	ALERT_INVALID_MIME_EXT	
ALERT_MALWARE_HASH_REGISTRY_MATCH	ALERT_MALICIOUS_URL	
ALERT_FAILED_PASSWORD	ALERT_CLEAR_HISTORY	
login	ALERT_ANOMALOUS_HOST	
ALERT_ILLEGAL_USER_ACTIVITY	ALERT_MULTIPLE_LOGIN	ALERT_WATCHED_COUNTRY_LOGIN

Comparison of DFG vs. baseline

Past attacks				
Method	*TP*	*TN*	*FP*	*FN*
Critical alert (KNOWN MALWARE)	8.95	1	0	91.05
Anomalous Alert (ANOMALOUS_HOST)	9.8	96.0	4.0	90.2
Dynamic Factor Graph	***74.2***	***98.5***	***1.5***	***25.8***
Replayed attacks				
Method	*TP*	*TN*	*FP*	*FN*
Critical alert (CVE-2017-5638)	8.4	99.55	0.45	91.3
Anomalous alert (RDP bad cookie)	0	99.96	0.037	1
Dynamic Factor Graph	***91.6***	***99.11***	***0.9***	***8.4***
Live traffic				
Method	*TP*	*TN*	*FP*	*FN*
Critical alert (Root Privilege)	N/A	93.1	6.9	N/A
Anomalous alert (SSH Bad client)	N/A	98.77	1.23	N/A
Dynamic Factor Graph	***N/A***	***0.993***	***0.007***	***N/A***

Table IV: Summary of results on critical and anomalous alerts observed in past, replayed, and live traffic attacks. TP: True Positive. TN: True Negative. FP: False Positive FN: False Negative

Summary of Dynamic Factor Graphs for Attack Preemption

Challenges

- Overwhelming noise and alert volumes
- Relying on late-state critical alerts
- Necessity of alert sequence detection

State of the Art

- Atomic alert-based system
- Black Hole Routing
- Deep generative models

Our Contributions

- **Characterization** of alert sequences in real-world, national scale HPC providing GPU resources
- **Insights:**

(1) Critical alerts arriving late

(2) Alert sequences length from two to four for detection

(3) Similarity of attacks (95% show 33% similar sequences)

(4) Timings of recurrent alerts matter.

- **Dynamic Factor Graphs:**

Uncertainty, conditional probabilities, and evolution of alerts

Future Work

- **Providing data on emerging vulnerabilities**

Accelerator drivers (0-day)

Provenance of AI models (data, code)

Quantum-resistant cryptography

MCP honeypot

- **Focus on attack recovery**

Automated integration with Black Hole Router

Minimize workload disruptions via checkpointing

- **New research security policies on AI/HPC**

Acknowledgements

National Science Foundation

Measurements [*USENIX NSDI, USENIX Security**]

400Gbps link

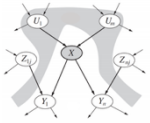
Optical tap

Network/host observability

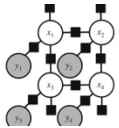
Host logs 100,000 notices/day

Modeling

Reinforcement Learning + POMDP



Bayesian Networks



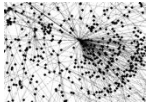
Markov Random Fields

Dynamic Factor Graphs

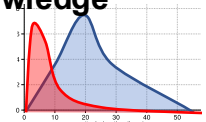
[AAAI-Workshop**]

[IEEE-CSR**]

Domain knowledge



Network Topology



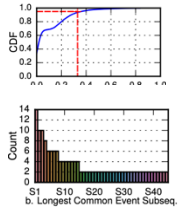
Conditional Probabilities

Alert Patterns

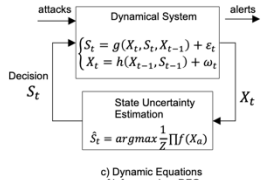
Expert' insights

* First author
** Collaboration
$ Grant

Data-driven Attacks



Formulation FFs [*TDSC-Submission]



Semi-automated Response

Dependable HPC systems

Stopped Ransomware family

Found Hidden Attacks

Novel Threats (PQC)

[*Secure HPC @ Supercomputing]

[NSF CC* $]

Technical Contributions

Results & Impact