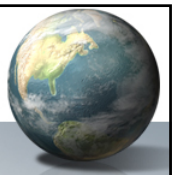
Early Intrusion Projection and Impact Assessment for Cyber Situational Awareness

Shanchieh Jay Yang[1]
Daniel Fava[1], Brian Argauer[1],
Jared Holsopple[2], Moises Sudit[2]

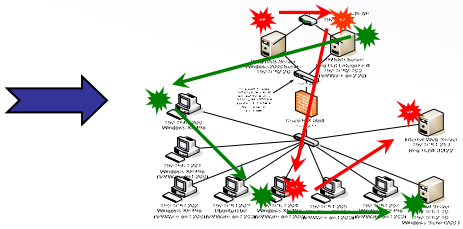
[1] Computer Engineering, RIT
[2] CMIF, SUNY Buffalo

Assessing Cyber Attacks



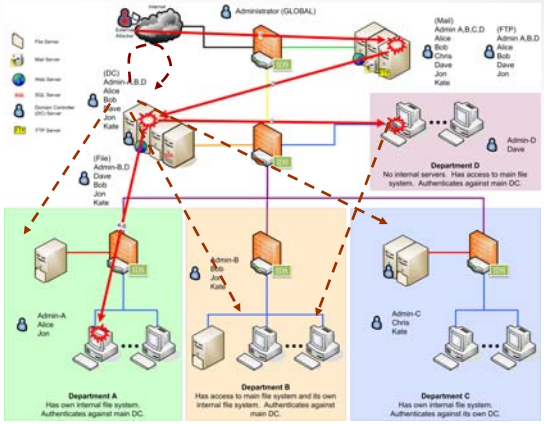
Q: What to do with overwhelming intrusion alerts?

- Alert aggregator/correlator forms attack tracks
- Estimate impact of observed attack tracks
  - Rule/scenario based [Porras & Fong '02], [Valeur, etal. '04]
- Project attack actions
  - Matching attack plans [Qin,Lee'04], [Wang, etal. '07]

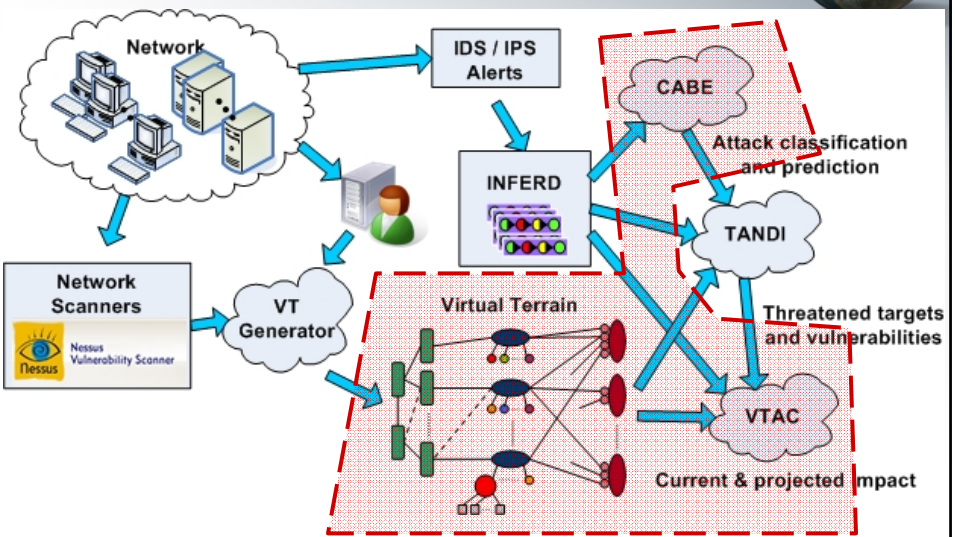


Current Impact vs. Future Threat

- Current: potential damage caused by observed attack tracks
- Future: anticipated moves (actions, targets, & time?) of attackers and the associated impact



Overall System

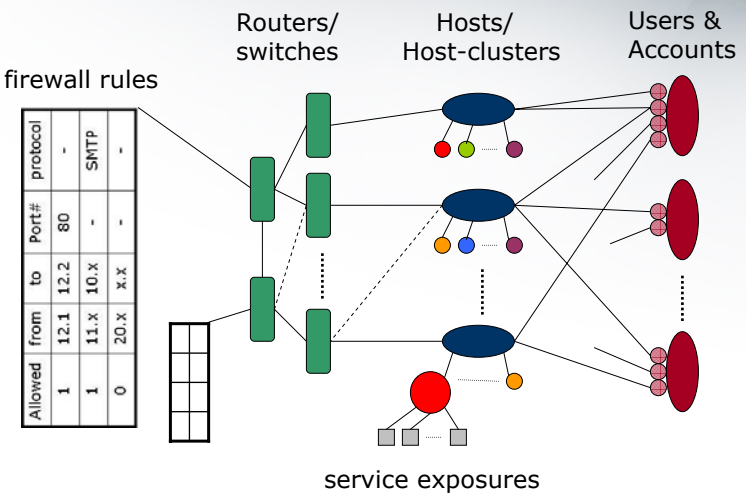


Cyber Context Needed

- Network information
  - Firewall / permission rules (analyst)
  - Users, accounts, and privileges (analyst)
  - Host, service, and user criticalities (analyst)
  - Local and remote services (scanner)
  - Mapping from services to vulnerabilities (scanner + databases)
  - Physical and logical subnet connectivity (scanner + analyst)
  - Exposure damage scores (common scoring system)
- Attack data
  - Network and host IDS alerts (correlation ground truth)
  - IDS locations (synchronized?)
  - Attack action (not alerts) ground truth?
  - Damage caused by attack actions (services and/or accounts compromised)

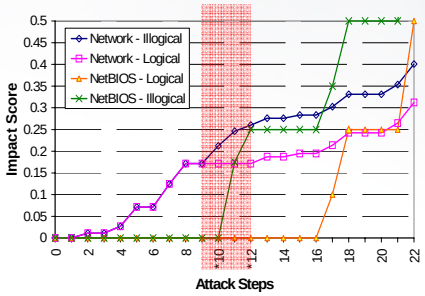
Virtual Terrain

- A graph-based model for projection & impact assessment



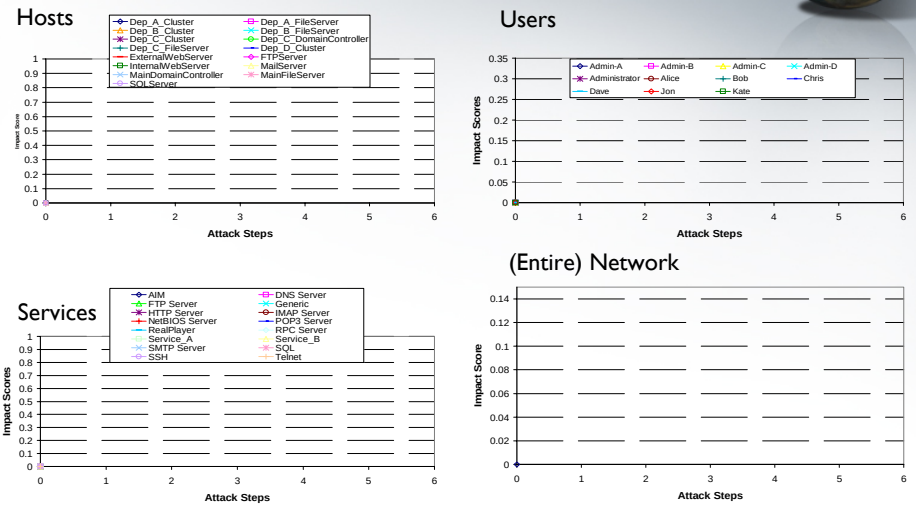
VTAC for Impact Assessment

- VTAC (Virtual Terrain assisted impact Assessment for Cyber attacks)
  - Graph-based VT reduce the complexity of defining attack prerequisite-consequence rules
  - Identify illogical steps, potentially due to
    - error in VT model, missing alerts, zero-day attacks, coordinated attacks
  - Ranks network entities for each attack track
  - Ranks attack tracks w.r.t. different impact scores



Step 0

Step	Description	Source IP	Dest IP	Port/Prot	Alert Signature
1	Compromise FTP Server	129.21.168.101	192.168.1.4	21/tcp	WEB-MISC /home/ftp access
2	Step to Internal Web Server	192.168.1.4	192.168.3.3	80/tcp	WEB-IIS .asa HTTP header buffer overflow attempt
3	Attack Department D Cluster	192.168.3.3	192.168.4.100	23/tcp	TELNET bsd telnet exploit response
4	Ping Dep A Cluster	192.168.3.3	192.168.11.103	456/icmp	ICMP PING Microsoft Windows
5	Scan Dep A Cluster	192.168.3.3	192.168.11.103	22/tcp	SCAN SSH Version map attempt
6	Attack Dep A Cluster	192.168.3.3	192.168.11.103	22/tcp	EXPLOIT ssh CRC32 overflow

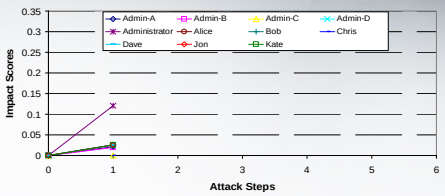


Step 1

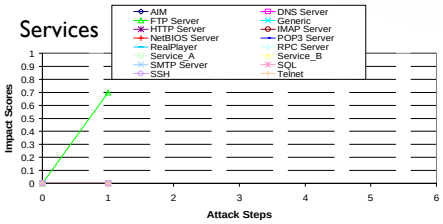
Step	Description	Source IP	Dest IP	Port/Prot	Alert Signature
1	Compromise FTP Server	192.21.168.101	192.168.1.4	21/tcp	WEB-MISC /home/ftp access
2	Step to Internal Web Server	192.168.1.4	192.168.3.3	80/tcp	WEB-IIS .asa HTTP header buffer overflow attempt
3	Attack Department D Cluster	192.168.3.3	192.168.4.100	23/tcp	TELNET bsd telnet exploit response
4	Ping Dep A Cluster	192.168.3.3	192.168.11.103	456/icmp	ICMP PING Microsoft Windows
5	Scan Dep A Cluster	192.168.3.3	192.168.11.103	22/tcp	SCAN SSH Version map attempt
6	Attack Dep A Cluster	192.168.3.3	192.168.11.103	22/tcp	EXPLOIT ssh CRC32 overflow

Hosts

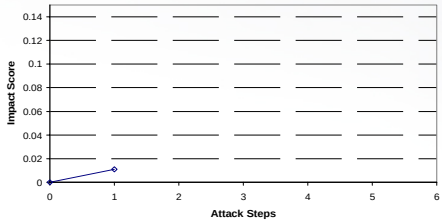
Users



Services



(Entire) Network

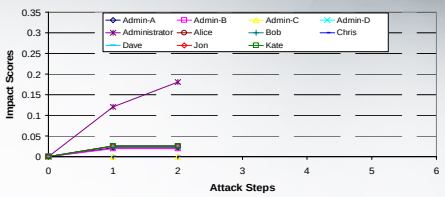


Step 2

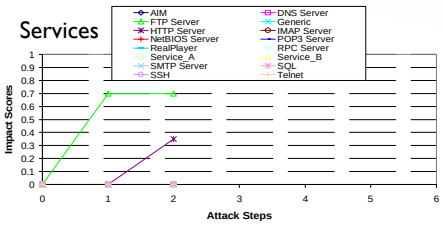
Step	Description	Source IP	Dest IP	Port/Prot	Alert Signature
1	Compromise FTP Server	192.21.168.101	192.168.1.4	21/tcp	WEB-MISC /home/ftp access
2	Step to Internal Web Server	192.168.1.4	192.168.3.3	80/tcp	WEB-IIS .asa HTTP header buffer overflow attempt
3	Attack Department D Cluster	192.168.3.3	192.168.4.100	23/tcp	TELNET bsd telnet exploit response
4	Ping Dep A Cluster	192.168.3.3	192.168.11.103	456/icmp	ICMP PING Microsoft Windows
5	Scan Dep A Cluster	192.168.3.3	192.168.11.103	22/tcp	SCAN SSH Version map attempt
6	Attack Dep A Cluster	192.168.3.3	192.168.11.103	22/tcp	EXPLOIT ssh CRC32 overflow

Hosts

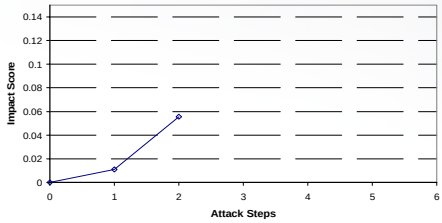
Users



Services



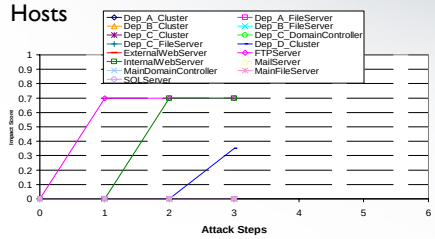
(Entire) Network



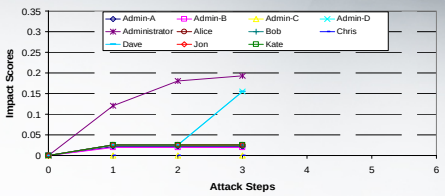
Step 3

Step	Description	Source IP	Dest IP	Port/Prot	Alert Signature
1	Compromise FTP Server	129.21.168.101	192.168.1.4	21/tcp	WEB-MISC /home/ftp access
2	Step to Internal Web Server	192.168.1.4	192.168.3.3	80/tcp	WEB-IIS .asa HTTP header buffer overflow attempt
3	Attack Department D Cluster	192.168.3.3	192.168.4.100	23/tcp	TELNET bsd telnet exploit response
4	Ping Dep A Cluster	192.168.3.3	192.168.11.103	456/icmp	ICMP PING Microsoft Windows
5	Scan Dep A Cluster	192.168.3.3	192.168.11.103	22/tcp	SCAN SSH Version map attempt
6	Attack Dep A Cluster	192.168.3.3	192.168.11.103	22/tcp	EXPLOIT ssh CRC32 overflow

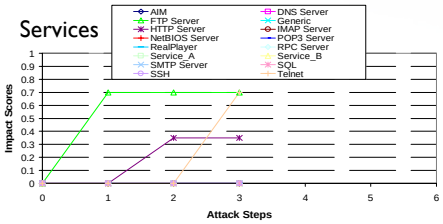
Hosts



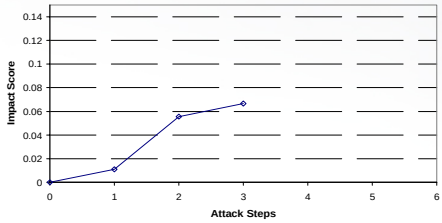
Users



Services



(Entire) Network

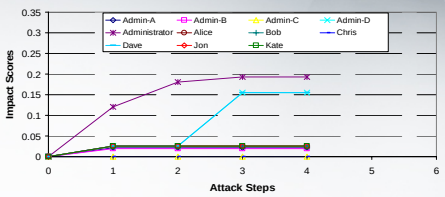


Step 4

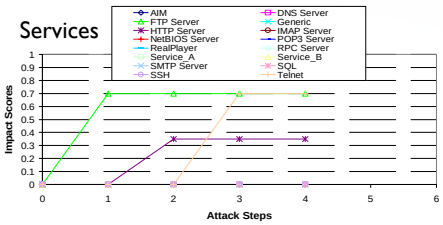
Step	Description	Source IP	Dest IP	Port/Prot	Alert Signature
1	Compromise FTP Server	129.21.168.101	192.168.1.4	21/tcp	WEB-MISC /home/ftp access
2	Step to Internal Web Server	192.168.1.4	192.168.3.3	80/tcp	WEB-IIS .asa HTTP header buffer overflow attempt
3	Attack Department D Cluster	192.168.3.3	192.168.4.100	23/tcp	TELNET bsd telnet exploit response
4	Ping Dep A Cluster	192.168.3.3	192.168.11.103	456/icmp	ICMP PING Microsoft Windows
5	Scan Dep A Cluster	192.168.3.3	192.168.11.103	22/tcp	SCAN SSH Version map attempt
6	Attack Dep A Cluster	192.168.3.3	192.168.11.103	22/tcp	EXPLOIT ssh CRC32 overflow

Hosts

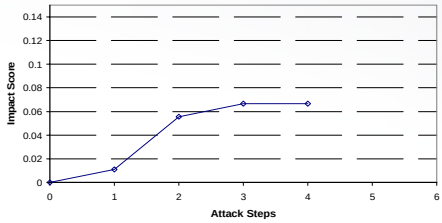
Users



Services



(Entire) Network

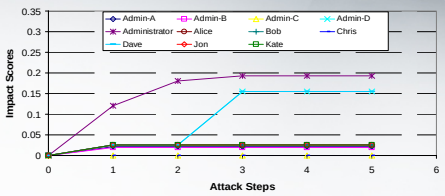


Step 5

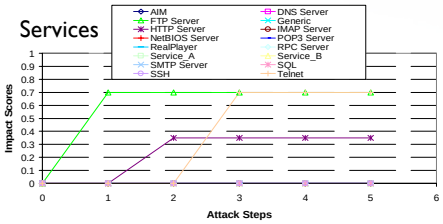
Step	Description	Source IP	Dest IP	Port/Prot	Alert Signature
1	Compromise FTP Server	129.21.168.101	192.168.1.4	21/tcp	WEB-MISC /home/ftp access
2	Step to Internal Web Server	192.168.1.4	192.168.3.3	80/tcp	WEB-IIS .asa HTTP header buffer overflow attempt
3	Attack Department D Cluster	192.168.3.3	192.168.4.100	23/tcp	TELNET bsd telnet exploit response
4	Ping Dep A Cluster	192.168.3.3	192.168.11.103	456/icmp	ICMP PING Microsoft Windows
5	Scan Dep A Cluster	192.168.3.3	192.168.11.103	22/tcp	SCAN SSH Version map attempt
6	Attack Dep A Cluster	192.168.3.3	192.168.11.103	22/tcp	EXPLOIT ssh CRC32 overflow

Hosts

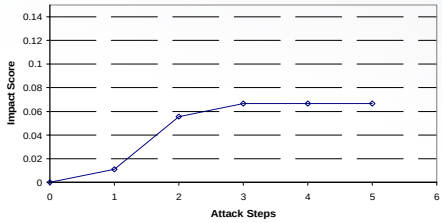
Users



Services



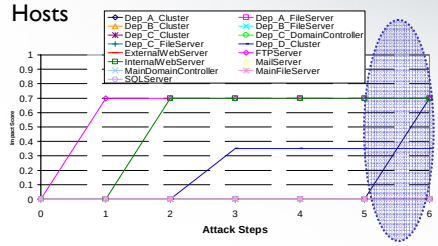
(Entire) Network



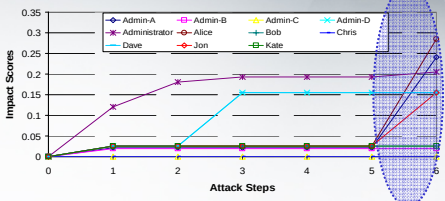
Projection

Step	Description	Source IP	Dest IP	Port/Prot	Alert Signature
1	Compromise FTP Server	129.21.168.101	192.168.1.4	21/tcp	WEB-MISC /home/ftp access
2	Step to Internal Web Server	192.168.1.4	192.168.3.3	80/tcp	WEB-IIS .asa HTTP header buffer overflow attempt
3	Attack Department D Cluster	192.168.3.3	192.168.4.100	23/tcp	TELNET bsd telnet exploit response
4	Ping Dep A Cluster	192.168.3.3	192.168.11.103	456/icmp	ICMP PING Microsoft Windows
5	Scan Dep A Cluster	192.168.3.3	192.168.11.103	22/tcp	SCAN SSH Version map attempt
6	Attack Dep A Cluster	192.168.3.3	192.168.11.103	22/tcp	EXPLOIT ssh CRC32 overflow

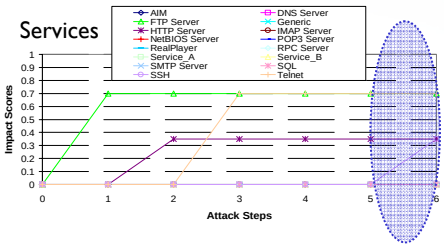
Hosts



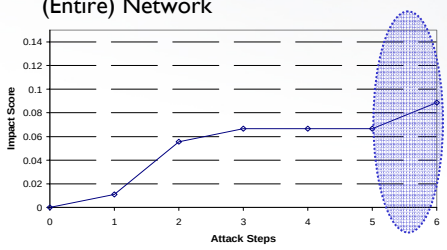
Users



Services

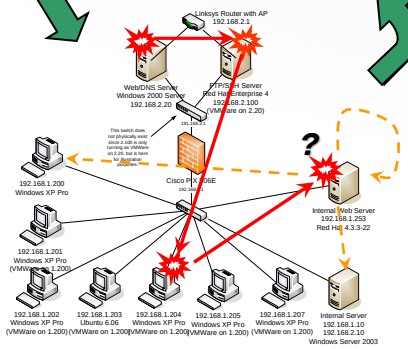


(Entire) Network



Projection: Terrain vs. Behavior

Behavior Models:
Behavior extraction using Variable Length Markov Model, Hidden Markov Model



Virtual Cyber Terrain:
Contextual reasoning based on logical connectivity, service, vulnerability, impact

Behavior analysis – how?

- Expert developed behavior model
  - E.g., guidance template, Bayesian Network
  - Diverse SME opinions (knowledge elicitation?)
  - Costly to maintain and update
- Attack tracks → time-stamp ordered sequences of symbols
- Context-based model
  - Adaptive Bayesian Network [Qin,Lee'04], Data Mining [Li etal.'07]
  - 0^{th}, 1^{st}, 2^{nd}, 3^{rd} order Markov Model
  - Variable-length Markov Model (VLMM)
  - Universal Predictor [Jacquet etal '02]
  - *Q*: What should be the context?
- State-based model
  - Hidden Markov Model (feasible?)

Alert Prediction Example

- Alerts Generated by Attack Actions

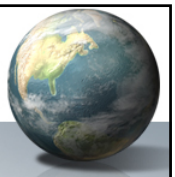
1. K (http_inspect) Oversize Request-URI Directory
2. F (http_inspect) Bare Byte Unicode Encoding
3. A ICMP PING NMAP
4. H ICMP L3retriever Ping
5. J WEB-MISC Invalid HTTP Version String
6. J WEB-MISC Invalid HTTP Version String
7. A ICMP PING NMAP
8. H ICMP L3retriever Ping
9. H ICMP L3retriever Ping
10. I NETBIOS SMB-DS IPC$ unicode share access
11. A ICMP PING NMAP
12. H ICMP L3retriever Ping

- Predicted Alerts

1. (no prediction)
2. J WEB-MISC Invalid HTTP Version String
3. J WEB-MISC Invalid HTTP Version String
4. H ICMP L3retriever Ping
5. A ICMP PING NMAP
6. J WEB-MISC Invalid HTTP Version String
7. J WEB-MISC Invalid HTTP Version String
8. H ICMP L3retriever Ping
9. F (http_inspect) Bare Byte Unicode Encoding
10. I NETBIOS SMB-DS IPC$ unicode share access
11. A ICMP PING NMAP
12. H ICMP L3retriever Ping

Bottom Line: Is there a reliable pattern (in what context) to extract for prediction?

Translating Alerts



- <Alert>
  - <Description>ICMP PING NMAP</Description>
  - <Dest_IP>100.20.0.0</Dest_IP>
  - <Category>Recon_Scanning</Category>
- </Alert>
- <Alert>
  - <Description>SCAN SOCKS Proxy attempt</Description>
  - <Dest_IP>100.10.0.1</Dest_IP>
  - <Category>Recon_Scanning</Category>
- </Alert>
- <Alert>
  - <Description>WEB-IIS nsiislog.dll access</Description>
  - <Dest_IP>100.20.0.0</Dest_IP>
  - <Category>Intrusion_Root</Category>
- </Alert>

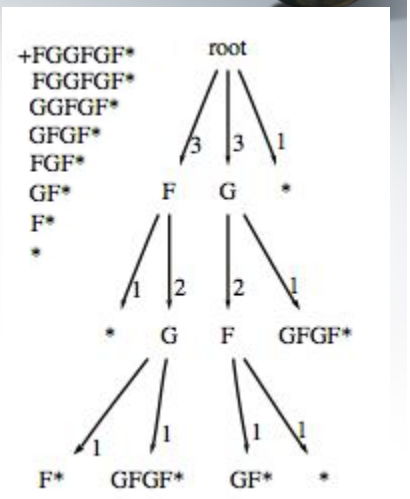
Category & target IP (Ω_t): AaB

Description (Ω_d): ABC

Category (Ω_c): AAB

Suffix Tree and Prediction

- +FGGFGF*
  - +: start of attack track
  - F: WEB-IIS nsiislog.dll access
  - G: WEB-MISC Invalid HTTP Version String
  - *: end of attack track
- What follows +GF?
  - -1th order: P=1/3
  - 0th order: P{G}=P{F}=3/7, P{*}=1/7
  - 1st order: P{G|F} = 2/3, P{*|F} = 1/3
  - 2nd order: P{G|GF} = 1/2, P{*|GF} = 1/2
  - VLMM – blending the estimates



Suffix tree from historical data

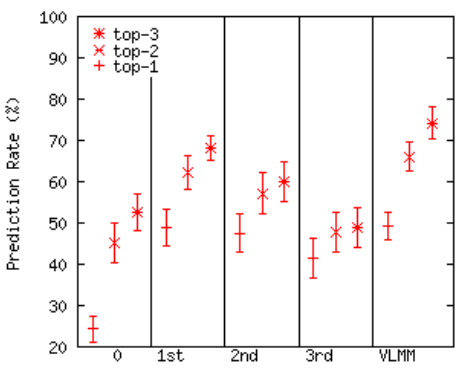
- Historical attack sequences builds suffix tree
- Suffix tree embeds patterns exhibited in finite-contexts
- Each unfolding attack sequence matches part of suffix tree for prediction

Experiment Setup

- Ground truth data generated via scripted attacks on a VMWare network
- A total of 1,113 attack sequences composed of 4,723 alerts after Δt=1 filtering [Valuer'04]
- 10 independent runs with random 50-50 splits of training vs. test data
- Alphabet choices:
  - Specific attack method (Ω_d)
  - Category of attack method (Ω_c)
  - Category + target IP (Ω_t)
- Top-*k* prediction rate (*k*=1, 2, 3):
  - % of correct prediction falls in the top-*k* choices

0 to 3rd Order and VLMM (Ω_d)

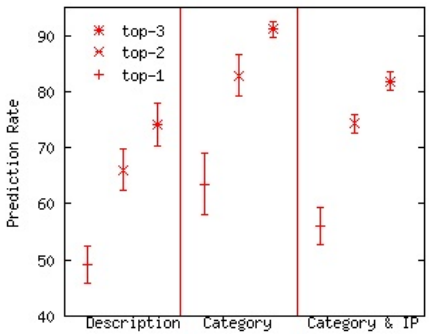
- Dominance of 1st order prediction
- VLMM combines n-order and offers better predictions
- Top 3 actions:
  - ICMP PING NMAP (43%), WEB-MISC Invalid HTTP Version String (22.4%), (http inspect) BARE BYTE UNICODE ENCODING (9.0%)
  - ICMP PING NMAP followed by ICMP PING NMAP 87.7% of the time
- Predicts better for repeating actions? Blending with longer context helps for predicting transitions?



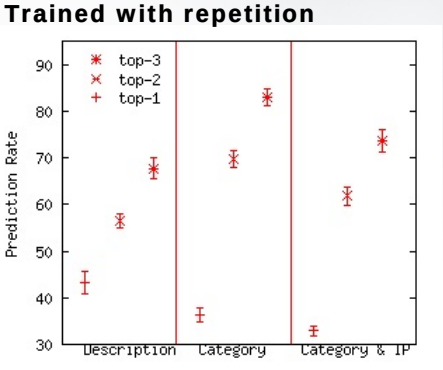
Prediction rate for transitions

- Predicting transitions will be better off by training with data sets with no repetition
- Predicting attack category is easier and more reasonable than predicting specific attack method

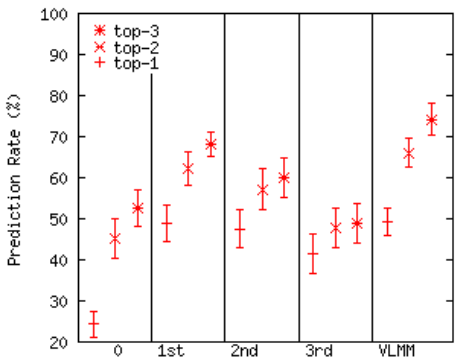
Trained with no repetition

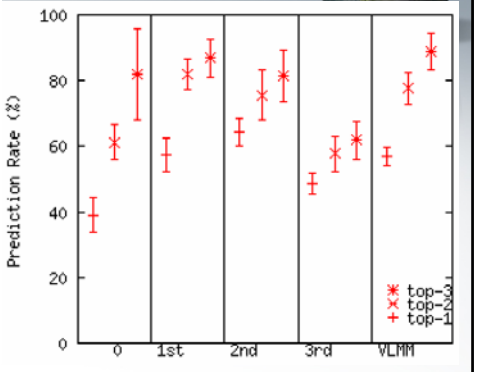


Trained with repetition



Ω_c better than Ω_d

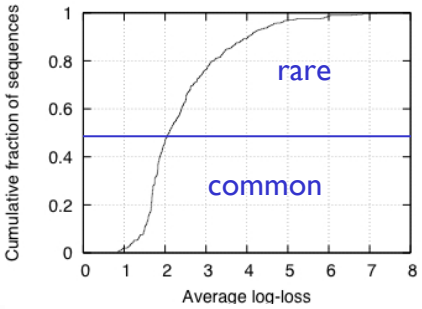


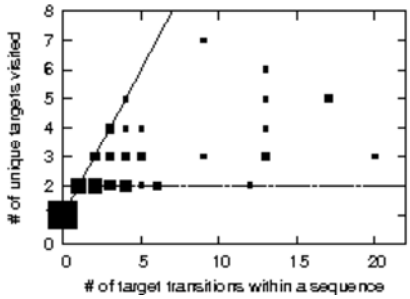


- Coarser granularity yield better prediction rate – as expected
- Network independent prediction – analyst may only want to know prediction at a coarser granularity

Classification?

- Can we categorize cyber attack types (with no ground truth)?
- Average Log-loss:
  - Rarity of attack sequence
  - Threshold=2.0 (Ω_c, no repetition)
    - 0.83 vs. 0.69 prediction rates
- # target trans vs. # targets visited:
  - Agility of attack
  - Most targets suffered 2 scans
  - Most popular targets: 1,735 and 814 out of a total of 4,723
  - Are more agile attacks harder to predict?





Conclusion

- Proactive impact assessment and projection of cyber attacks!!
- Graph-based VT defines dynamic relationships between network entities
  - Automatic update is not an easy task
- VTAC determines attack's impact to network elements
  - How to validate its performance?
- Behavior-based attack prediction
  - A new theoretical and real-world problem
    - Diverse, changing, and noisy behavior