Using AI to Enhance Cybersecurity

V.S. Subrahmanian
Northwestern Security & AI Lab
vss@northwestern.edu

June 6 2023
Indo-Pacific GeoIntelligence

Talk Outline

1. Northwestern Cyber Early Warning System: NCEWS
2. Estimate Malware Spread
3. Deter IP Theft
4. Deepfakes
5. NATO Locked Shields Research

Cyber Early Warning Systems

New Vulnerability Discovered

Will it be used in a real-world exploit?

If yes, how far into the future?

How severe will it be?

- What percentage of discovered vulnerabilities are used in an attack?
- What percentage of known vulnerabilities are used in an attack before NIST releases their assessment (CVSS) ?

The faster these questions can be answered, the safer the enterprise will be

Cyber Early Warning Systems

New Vulnerability Discovered

Will it be used in a real-world exploit?
> 95 % F1-score

If yes, how far into the future?
+/- 11.9 days error

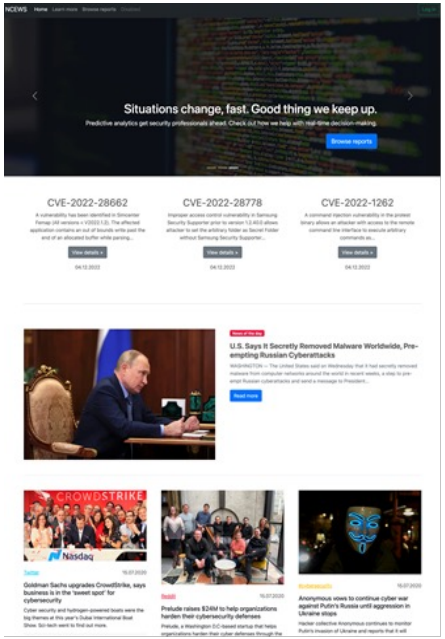
How severe will it be?
+/- 1.25 error on a 10 point scale

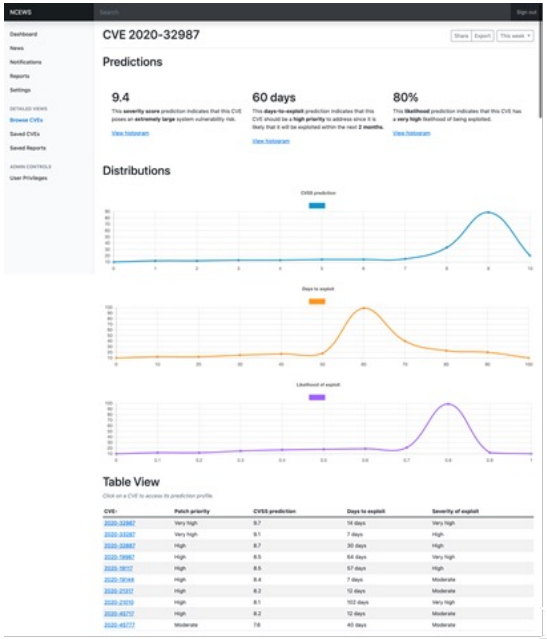
What percentage of discovered vulnerabilities are used in an attack? **9.2%**

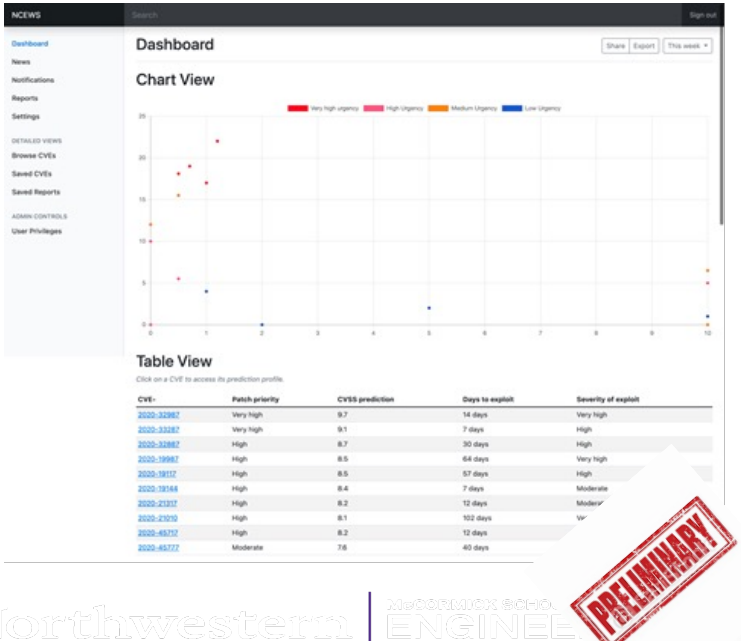
What percentage of known vulnerabilities are used in an attack before NIST releases their assessment (CVSS) ? **49.46%**

The faster these questions can be answered, the safer the enterprise will be

NCEWS Screenshots







Estimating Malware Spread

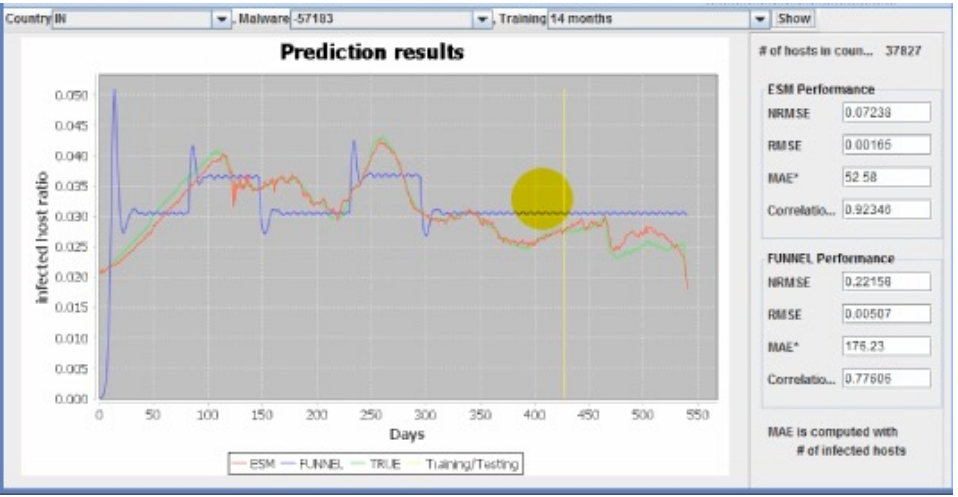
New Malware Discovered

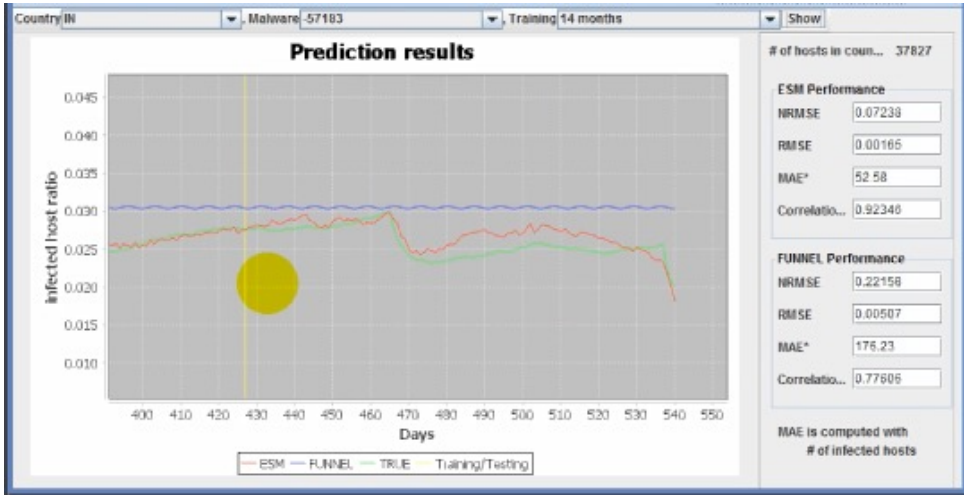
What percentage of hosts in my network will be hit?

What percentage of high asset value hosts will be hit?

How best can we mitigate this spread?

Estimating Malware Spread





Using data from Symantec

Deterring IP Theft with Fakes

Sentences from Restaurant Reviews

1. decent black coffee and good service
2. authentic tasting sauce, fresh crust

Sentences from Movie Reviews

3. Russell Crowe's direction is far too predictable
4. the movie just feels dry and generic and very narrow all around

Scientific Paper Titles

5. sequential depth quantization method for large-scale reconstruction
6. parsing and ensemble of deep convolutional neural network models

Which are real which are fake?

Generating Fakes to Deter IP Theft

Sentences from Restaurant Reviews

1. decent black coffee and good service
2. authentic tasting sauce, fresh crust

Sentences from Movie Reviews

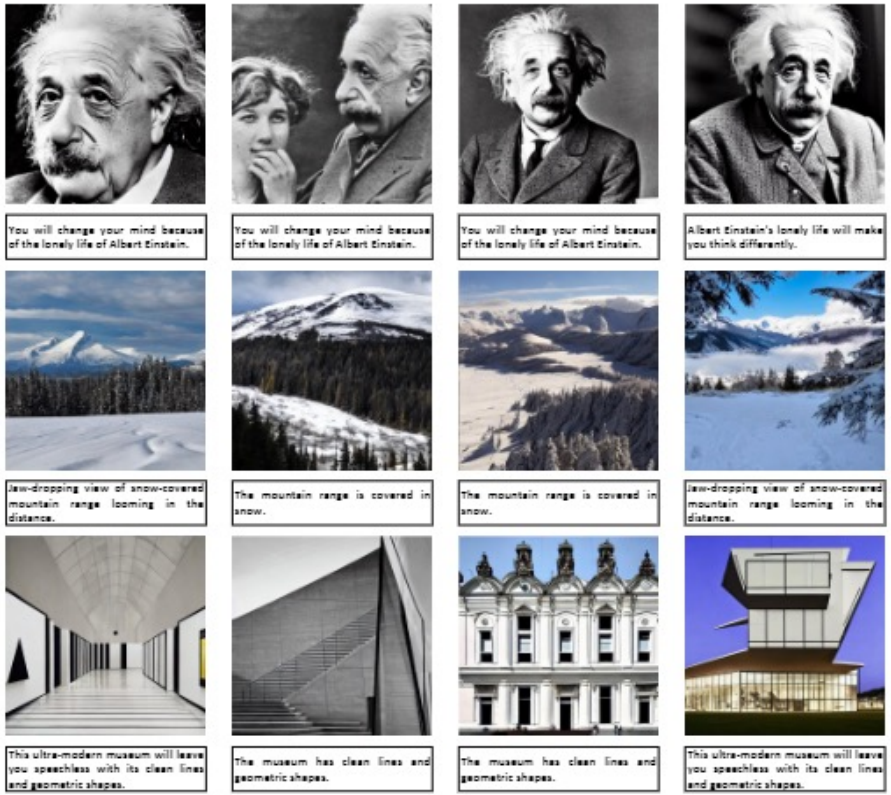
3. Russell Crowe's direction is far too predictable
4. the movie just feels dry and generic and very narrow all around

Scientific Paper Titles

5. sequential depth quantization method for large-scale reconstruction
6. parsing and ensemble of deep convolutional neural network models

Sentences in red are fake

Generating Multimodal Fakes



Extended to handle data theft as well

Identifying Suspicious Sessions in the NATO Locked Shields Exercise

- Huge cybersecurity exercise (32 countries, 2000+ participants in 2022)
- 1000+ servers on the Locked Shields network
- Over 14M sessions of data. ~350K are malicious.
- Goal: Find suspicious sessions. Fast.

Joint work with Netherlands Defence & TU Delft

Some Prediction Results

- Predictions after
  - 4 hours
  - 8 hours
  - 16 hours
  - 32 hours
- Our **suspicion score** approach outperforms existing approaches in all settings

After 4 hours

Approach	AUC	Precision	Recall	F1-Score
SUS	**1**	1	1	**1**
Swiss[1]	0.999	0.604	1	0.753
Jan-Klein[2]	0.859	0.404	0.719	0.517

After 8 hours

Approach	AUC	Precision	Recall	F1-Score
SUS	**0.99**	0.9	0.98	**0.938**
Swiss[1]	0.838	0.59	0.676	0.633
Jan-Klein[2]	0.75	0.386	0.505	0.438

After 16 hours

Approach	AUC	Precision	Recall	F1-Score
SUS	**0.995**	0.699	0.99	**0.819**
Swiss[1]	0.874	0.437	0.75	0.55
Jan-Klein[2]	0.904	0.705	0.809	0.754

Conclusion

- AI is already playing a major role in cyber-defense
  - Malware analysis and detection
  - Corporate risk posture
  - Intellectual property protection, data breaches
- The role of AI is poised to grow in coming years.
  - Generating and detecting phishing messages and posts
  - Generating and detecting infected URLs
  - Generating and detecting attack graphs
  - Generating and detecting malware

Contact Information

V.S. Subrahmanian
Walter P. Murphy Professor of Computer Science
Buffett Faculty Fellow – Buffett Institute of Global Affairs
Northwestern University
1800 Sherman Ave, Suite 3-000
Evanston, IL 60201.
vss@northwestern.edu
https://vssubrah.github.io/