Cybersecurity threats – Can we predict them?

Cyber attacks on computers and associated infrastructure are becoming increasingly sophisticated and it has become evident that it is no longer effective to deal with cyber attacks after they occur. A new model is needed based on the prediction and prevention of potential attacks. Professor Shanchieh (Jay) Yang, a researcher and Department Head for the Department of Computer Engineering at the Rochester Institute of Technology and collaborators are developing a system that will identify and predict critical attack behaviours, simulate real and theoretical cyber attacks scenarios and forecast attacks before they happen.

Cyber attacks on computers and associated equipment and other infrastructure have become increasingly sophisticated in recent years and are also increasing in frequency. The motivation behind most cyber attacks has shifted from thrill seeking or notoriety to the pursuit of profit or political gain. Cybercrime has enormous associated costs for large corporations and government institutions, with US organisations having the highest average cost of cybercrime ($17.36 million), followed by Germany ($17.36 million) and the UK ($7.21 million). Attacks are often large-scale, multidimensional and involve a variety of techniques to sabotage computer systems or obtain valuable information. Large-scale cyber attacks include botnets, where many hosts perform similar actions, or can also consist of a team of colluding sources.

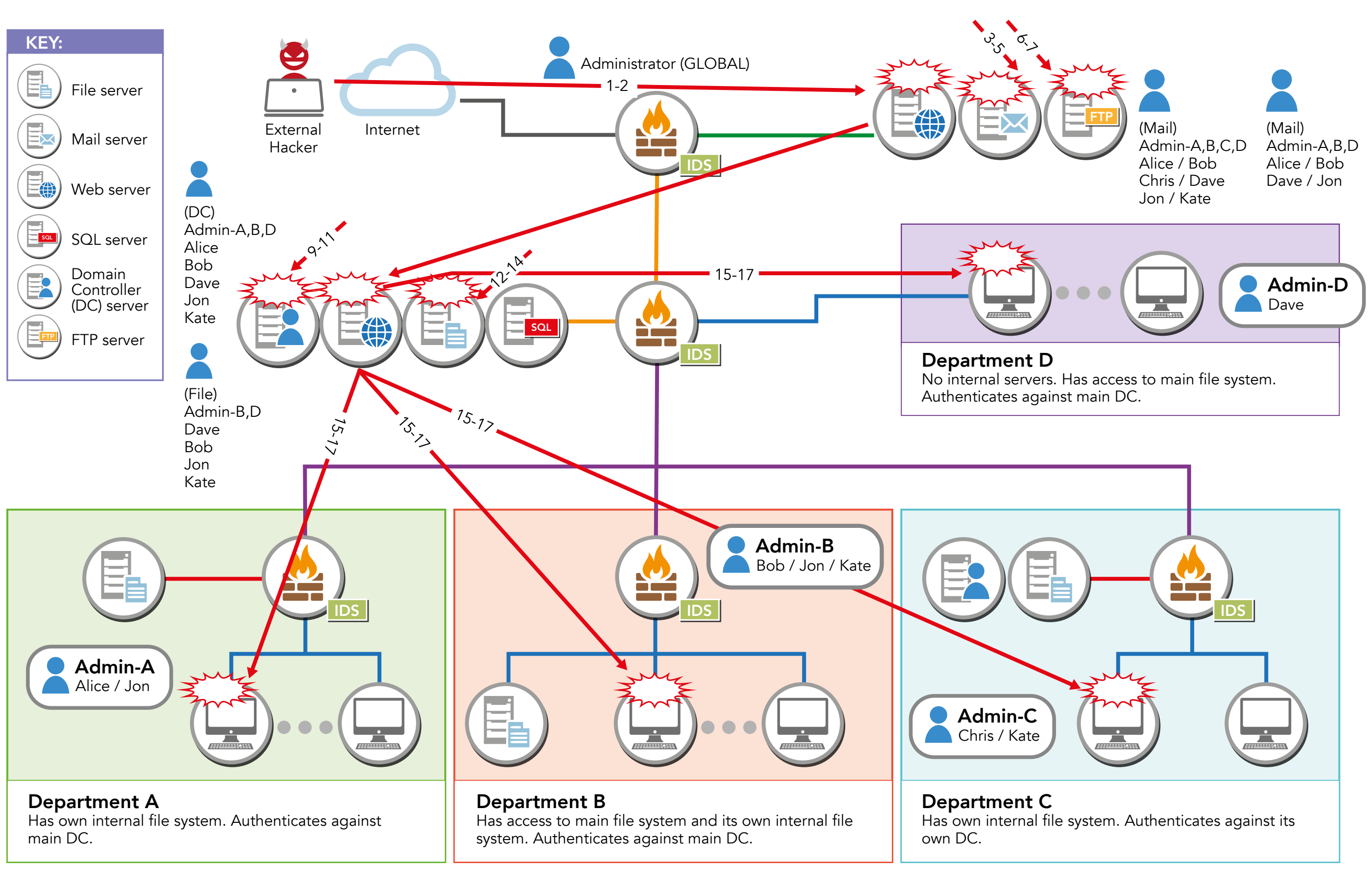
Current responses to cyber attacks are defensive or reactionary, meaning that the attacks are only removed and analysed after systems have been exploited. Some common defence strategies include intrusion detection and prevention tools such as antivirus software, the use of firewalls, and access controls such as passwords. However, these are costly, time intensive, and becoming increasingly ineffective as cyber attacks become larger, more coordinated and harder to detect.

Researchers are now seeking more effective ways to predict and plan for the multitude of cyber attacks which could happen in the future, as opposed to relying solely on defensive or removal measures. Professor Yang and his team at Rochester Institute of Technology are developing a novel predictive system based on a combination of algorithms, machine learning and criminology theory that will analyse critical attack behaviours, simulate cyber attack scenarios to understand real and theoretical attacks, and forecast when these attacks will occur.

CURRENT PREDICTION TOOLS ARE LIMITED

Various methods for the prediction of cyber attacks have been designed in the past, but with limited success. Prediction based on network vulnerabilities can be effective but requires up-to-date knowledge of the network, which is challenging. These methods anticipate future attacks by identifying malicious activities occurring in networks. However, the presence of decoy or obscured attacks can also cause errors and therefore not perfect. To discover sophisticated attacks buried in the overwhelming data requires equally advanced analytical and prediction techniques.

Attack graphs are one tool researchers have used to perform cyber threat prediction. Attack graphs show most, if not all the ways in which a hacker can exploit vulnerabilities to break into a network of a computer system and this data can be analysed to see where a systems' weaknesses lie. An alternative to the use of attack graph is the use of a Dynamic Bayesian Network (DBN), a type of statistical model which estimates probabilities over time, revealing any patterns present in the attacks. Bayesian networks are typically used to predict the ultimate goal of an attack, for example, whether the attacker will compromise an account and password.



An example of cyberattack progressing into the network through a plethora of techniques.

Another technique for threat prediction is one that estimates the capability, opportunity, and intent of the attacker (COI). This method has been used widely in military and intelligence communities for threat assessment. Capability predicts which services the attacker is likely to target based on what he or she has successfully exploited before. Opportunity investigates whether an attacker has insider information of the network and what kinds of safeguards the network has. Intent is the study of attacker motivation and social influence. Unfortunately, network attack projection via COI analysis is at its early stages and does not work well for attacks that constantly change the strategy.

> **Cyber threats evolve in a dynamic way. Therefore, they need a dynamic response.**

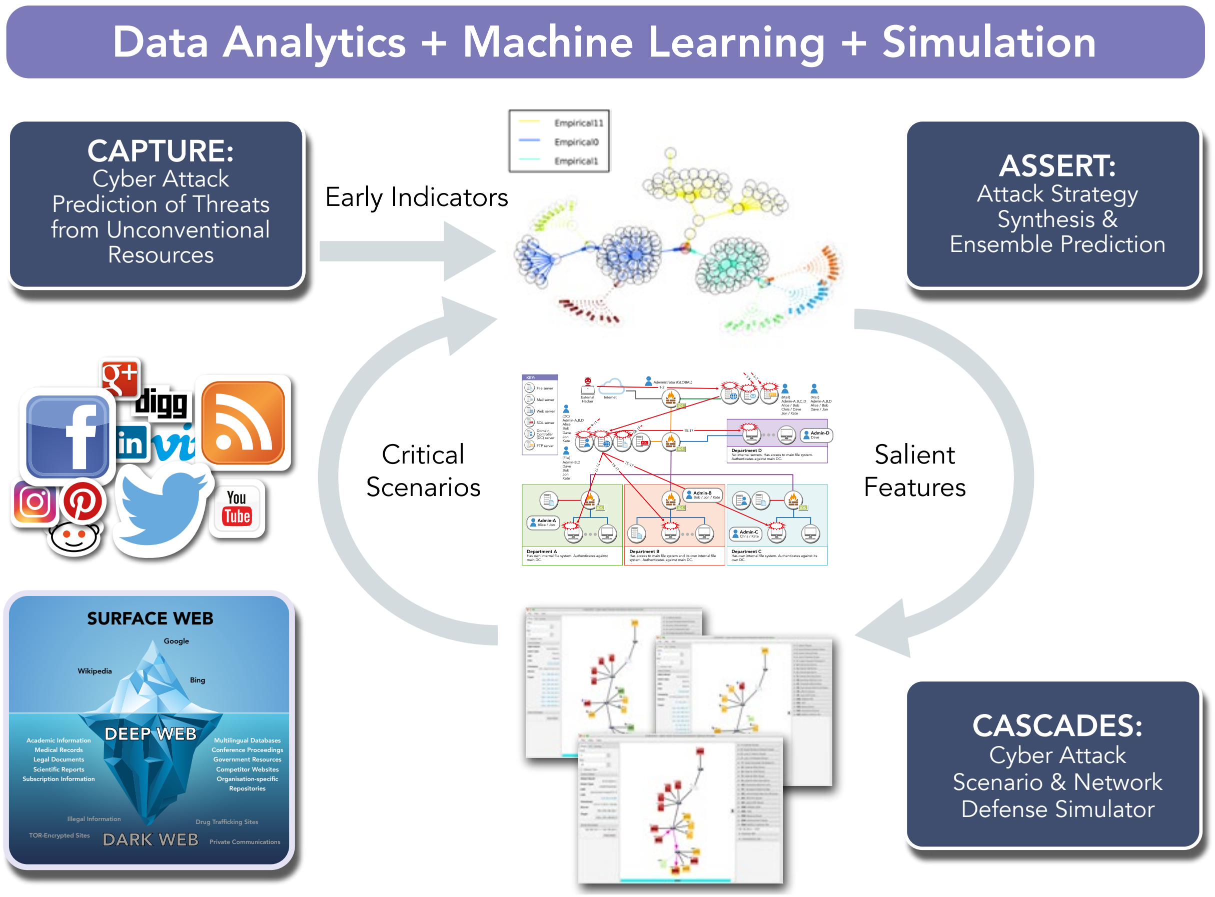
Recommendation systems, typically used for movie ranking and shopping sites, have also been used to predict which networks might be vulnerable based on the behaviour of malicious source Internet Protocol Addresses (IP). IP is a numerical label assigned to each device connected to a computer network.

The above cyber threat prediction systems have shown promises and limitations due to errors in intrusion detection, incomplete network information and attack obfuscation. Obfuscation techniques are those used to evade detection by making malicious code deliberately hard to understand, for example, the insertion of noise into malware to evade detection by the network. Combating against large-scale, coordinated attacks requires advances on various fronts, including intrusion detection, alert correlation (the collection of events generated from computer systems), attack characterisation, attack prediction, and host clustering.

When assessing the networks' security and risks, hacker behaviours also need to be taken into consideration, which can be a daunting task due to the vast number of potential known and unknown vulnerabilities in the network and choices an attacker could make to penetrate into a network.

COMBINING APPROACHES TO CYBER ATTACKS PREVENTION

Professor Yang's group is working on two projects, both funded by the National Science Foundation (NSF) in the US. The first project – Attack Strategy Synthesis and Ensemble Predictions of Threats (ASSERT), uses observable malicious activities occurring on networks to predict imminent attacks. With ASSERT, Dr Yang anticipates it will be possible to develop a strategy to differentiate ongoing malicious activities and respond to the upcoming critical threat before such an event occurs. ASSERT uses data sources found on networks such as intrusion detection system alerts and system logs, to recognise and predict imminent threats. The project integrates adaptive Bayesian learning, Clustering, and Information Theory-based divergence measures to generate and refine hypothetical attacks on computer networks. Meanwhile, the same data is fed to Long-Short-Term-Memory (LSTM) Network and Generative Adversarial Network (GAN), two deep learning


Data Analytics + Machine Learning + Simulation
CAPTURE:
Cyber Attack Prediction of Threats from Unconventional Resources
Early Indicators
ASSERT:
Attack Strategy Synthesis & Ensemble Prediction
Critical Scenarios
Salient Features
SURFACE WEB
DEEP WEB
DARK WEB
CASCADES:
Cyber Attack Scenario & Network Defense Simulator

Achieving predictive cyber defense through an ensemble of advances in data analytics, machine learning and simulation.

techniques that deals with sequential data, to characterise cyber attacks.

The second NSF funded project – Cyber Attack Scenario and Network Defence Simulator (CASCADES), aims at simulating cyber attack scenarios based on renewed criminology theory for cyber criminals. The scenarios enable what-if analysis and thus enable forecasting the most likely types of threats that could transpire in the future. CASCADES generates attack scenarios utilising Monte-Carlo and Importance Sampling and includes factors such as attacker COI and preference. A Monte Carlo simulation is a powerful statistical analysis tool that generates chance variables and exhibits random behaviours and has been used by professionals in many fields such as finance and engineering. Importance Sampling is a statistical technique for estimating the properties of a group of potential outcome probabilities. Simulations are run for multiple network configurations and multiple attacker types such as experts, amateurs, or random attackers. These simulation scenarios are exploratory and theory-based and can give researchers a potentially not-yet-explored understanding of adversary behaviours.

As the crime process unfolds, the interaction between the offender and target fluctuates, which determines whether the crime will stop or progress to the next stage. Data from real and hypothetical cyber attacks are fed to both ASSERT and CASCADES, creating a dynamic system where the two systems can feed data to each other as it is produced, enhancing the learning process. The CASCADES project can feed ASSERT simulated data and ASSERT can also provide data to CASCADES that can be used to guide simulation.

> **The internet will be more secure if we will be able to one day to predict or anticipate cyber attacks.**

THE FUTURE OF CYBER-SECURITY

Dr Yang states that the ASSERT and CASCADES projects are constantly evolving as more is learned about the ever-evolving cyber terrorism techniques and criminal behaviours of perpetrators. A prototype is still a few years away. However, cybersecurity is at a crucial turning point and computer systems globally are in dire need of a system based on comprehensive and anticipatory cyber threat analysis. Future research efforts in the field must focus on the development of cyber attack prediction systems that can anticipate critical scenarios and outcomes rather than relying on defensive solutions or focusing on damage control. Professor Yang's team leads the way in the development of sophisticated models that will be able to predict and forecast large-scale cyber threats before they take place and allow organisations effective solutions in the ongoing battle against cybercrime.

Behind the Research

Professor Shanchieh Jay Yang

E: Jay.Yang@rit.edu **T:** +1 585 475 6434 **W:** https://nsf.gov/awardsearch/showAward?AWD_ID=1526383&HistoricalAwards=false
W: www.nsf.gov/awardsearch/showAward?AWD_ID=1742789&HistoricalAwards=false **W:** www.rit.edu/news/story.php?id=53394
W: www.rit.edu/news/story.php?id=59040 **W:** www.rit.edu/news/story.php?id=63568 **W:** www.rit.edu/cybersecurity/
W: https://people.rit.edu/sjyeec

Research Objectives

These projects led by Professor Yang, investigate how to extract critical attack attributes, synthesise novel attack sequences, and reveal potential threats to critical assets in a timely manner. The projects use machine learning and simulation techniques to simultaneously identify emerging attack scenarios and observed events that could identify those attacks.

Detail

Shanchieh Jay Yang
PhD, ECE, the University of Texas at Austin
Professor & Department Head
Department of Computer Engineering
Rochester Institute of Technology
83 Lomb Memorial Drive, Bldg 09
Rochester, NY 14623-5603
Office: Bldg. 09, Room 3480
USA

Bio
Dr S. Jay Yang received his PhD in Electrical and Computer Engineering from the University of Texas at Austin. He is currently a Professor and the Department Head of the Department of Computer Engineering at Rochester Institute of Technology.

Funding
National Science Foundation (NSF)

Collaborators

- Dr Aunshul Rege, Criminal Justice Department, Temple University
- Dr Michael Kuhl, Industrial and System Engineering, Rochester Institute of Technology
- Mr Daryl Johnson, Computing Security, Rochester Institute of Technology
- Mr Bill Stackpole, Computing Security, Rochester Institute of Technology

References

www.infosecurity-magazine.com/next-gen-infosec/ai-future-cybersecurity/

https://people.rit.edu/~sjyeec/

www.palisade.com/risk/monte_carlo_simulation.asp

Haitao Du, Shanchieh Jay Yang. (2013). 'Sequential modeling for obfuscated network attack action sequences'. Communications and Network Security (CNS) 2013 IEEE Conference pp. 389-390.

Soorena Merat, Wahab Almuhtadi. 'Artificial intelligence application for improving cyber-security acquirement'. *Electrical and Computer Engineering (CCECE), 2015 IEEE 28th Canadian Conference*, 1445–1450.

S. Moskal, S. J. Yang, and M. Kuhl. (2018).'Cyber Threat Assessment via Attack Scenario Simulation using an Integrated Adversary and Network Modeling Approach'. *Journal of Defense Modeling and Simulation*, Vol. 15, No.1, pp.13-29.

J. Holsopple, S. J. Yang, and M. Sudit. (2015). 'Mission Impact Assessment for Cyber Warfare'. Book chapter in R. R. Yager, M. Z. Reformat, and N. Alajlan (Eds.), *Intelligent Methods for Cyber Warfare*, Springer, pp. 239-266.

H. Du, C. Wang, T. Zhang, S. J. Yang, J. Choi, and P. Liu. (2015). "Cyber Insider Mission Detection for Situation Awareness". Book chapter in R. R. Yager, M. Z. Reformat, and N. Alajlan (Eds.), *Intelligent Methods for Cyber Warfare*, Springer, pp. 201-218.

J. Holsopple, M. Sudit, and S. J. Yang. (2014)."Top-down Driven Cyber Impact Assessment". Book chapter in A. Kott, R. Erbacher, and C. Wang (Eds.), *Cyber Defense and Situational Awareness*, Springer, pp. 219-238.

S. J. Yang, H. Du, J. Holsopple, and M. Sudit. (2014). 'Attack Projection for Predictive Cyber Situation Awareness'. Book chapter in A. Kott, R. Erbacher, and C. Wang (Eds.), *Cyber Defense and Situational Awareness*, Springer, pp. 239-261.

Personal Response

How successful has your model been and how long will it take to be fully operational?

“ Using data collected through National Collegiate Penetration Testing Competition (http://nationalcptc.org/), ASSERT is able to separate common and unique cyber attack tactics, and predict previously unseen use of exploits towards a target or from a source. CASCADES has been shown to perform what-if analysis when a network misconfiguration is introduced, and demonstrate that more experienced hackers will take advantage of the misconfiguration but not as much for the novice hackers. Both systems are operational in a research setting and continue being enhanced. Professor Yang and his team are also looking for transition partners to achieve a broader impact for anticipatory cyber defence. ”