جامعة محمد الخامس بالرباط
Université Mohammed V de Rabat

École Nationale Supérieure d'Informatique et d'Analyse des Systèmes
Centre d'Études Doctorales en Sciences des Technologies de l'Information et de l'Ingénieur

AVIS DE SOUTENANCE DE THÈSE DE DOCTORAT

Monsieur HAMMOUCHI HICHAM

Soutiendra publiquement sa thèse de Doctorat en Informatique

Spécialité : Informatique

Le 09 Novembre 2023/2024 à 10h00 au Grand Amphi à l'ENSIAS de Rabat

Intitulé de la thèse

Socio-Technical Modeling for Proactive Cyber Attack Prediction

Devant le Jury composé de :

Président :

Pr. Abdellatif Kobbane, PES, ENSIAS, Université Mohammed V de Rabat

Directeur de thèse :

Pr. Houda Benbrahim, PES, ENSIAS, Université Mohammed V de Rabat

Co-Encadrant :

Pr. Mounir Ghogho, PES, Université Internationale de Rabat

Rapporteurs :

Pr. Abdellatif Kobbane, PES, ENSIAS, Université Mohammed V de Rabat

Pr. Jean-Yves Marion, PES, Université de Lorraine, France

Pr. Khalid Chougdali, PES, ENSA, Université Ibn Tofail, Kenitra

Examinateur :

Pr. Mohammed Boulmalf, PES, Université Internationale de Rabat

UNIVERSITE MOHAMMED V - RABAT
ENSIAS

Abstract: Cyberattacks are becoming increasingly complex as we become more reliant on technology. They can inflict increasing amounts of damage worldwide in ever-increasing ways. The increase in attack surfaces amplifies the risk of data leaks and the exposure of confidential data. In light of this complexity, researchers geared their focus to proactive and predictive systems that learn from historical incidents to predict the likelihood of occurrence of future attacks. However, most of prior work overlooked the contextual aspect and focused mainly on the technical aspects of attacks. These technical approaches, however necessary to build secure systems, should be reinforced with contextual knowledge in order to gain more against attackers' strategies.

In this thesis we address the problem of cyber attacks prediction and take a data-driven and a socio-technical modeling method inspired by socio-technical systems to account for social signals and contextual information when modeling cyber attacks. We leverage social data to design a rich and comprehensive socio-technical posture for organizations. This posture takes into account technical indicators, such as network logs and misconfigurations, as well as focused social media, mainly Twitter data, combined altogether to build predictive models to predict future severity levels and data breaches at corporate level. In this thesis, we aim to answer the central research question: “How can we model both technical and social cybersecurity aspects to predict future cyber attacks?” through several sub-questions that address nation-level and corporate-level attacks. By answering these questions, we make several contributions including: 1) we model the prediction of future attack rates severity as a socio-technical problem and build machine learning models for an accurate prediction. We focus on the severity rather than the actual rates to aid analysts get an informed picture of the present threats. 2) We develop STRisk where we design a socio-technical posture for organizations composed of technical network misconfigurations and a social posture inferred from social media to assess organizations security risk. Using STRisk we predict the likelihood of occurrence of data breaches at corporate-level in the US with an excellent accuracy. 3) We address the major issue of unreported data breaches and propose a data-driven technique to accurately detect them using the socio-technical posture.

This thesis has broadened the scope of cyber attacks prediction, and the presented techniques can serve as tools in the arsenal of system administrators and security analysts to better defend their assets, aid them strengthen the organizational security posture and mitigate the damage of future attacks. Indeed,

we demonstrate the significant added value of considering contextual and social signals to bring further insights into the security posture. Using the presented techniques, professionals can make informed decisions as the system are interpretable. Besides, although the proposed methods are presented in specific use-case, but can easily be generalized and adapted to other or similar use-cases and organizations.

Keywords: Cyber-attack prediction, Data breaches, Data-driven analysis, Machine learning, Predictive Systems, Socio-technical modeling.