Seminar at the Lab of Data Science, VIASM, August 18th, 2022

Title: Adversarial Attacks to Recommender Systems in Software Engineering

Speaker: Dr. Phuong Nguyen, University of L'Aquila (Italy)

Abstract: In recent years, we have witnessed a dramatic increase in the application of Machine Learning (ML) algorithms in several domains, including the development of recommender systems for software engineering (RSSE). Among others, library and API recommender systems have gained momentum as they became more successful at suggesting third-party libraries, API calls or code snippets. While these systems have proven to be effective in terms of prediction accuracy, there has been less attention for what concerns such recommenders' resilience against adversarial attempts. In fact, by crafting the recommenders' learning material, e.g., data from large open-source software (OSS) repositories, hostile users may succeed in injecting malicious data, putting at risk the software clients adopting RSSE.

In this seminar, we are going to present an empirical investigation of adversarial machine learning techniques and their possible influence on RSSE. The evaluation performed on different RSSE reveals a worrying outcome: all of them are not immune to malicious data. The obtained result triggers the need for effective countermeasures to protect recommender systems against hostile attacks disguised in training data.

The main topics presented in the seminar are as follows:

- Adversarial machine learning in Software Engineering.
- Possible countermeasures.

Short bio: Dr. Phuong Nguyen obtained a Ph.D. in Computer Science from Friedrich-Schiller-Universität Jena (Germany). He has worked as a research and teaching assistant at various universities in Vietnam. In 2014, Phuong was a postdoctoral researcher at Politecnico di Bari (Italy), working with recommender systems, Semantic Web, and Linked Data. After that, from August 2017 to January 2022, Phuong held a position as a postdoctoral researcher at Universit`a degli Studi dell'Aquila (Italy). Since February 2022, he has been a tenure track assistant professor at the same university, doing research in Software Engineering, Model-Driven Engineering, and Machine Learning. His research interests include Machine Learning developments in Software Engineering and Model-Driven Engineering with applications in computer networks, semantic web, recommender systems, and classification/clustering of modeling repositories. Phuong has worked on different European projects including CROSSMINER and TYPHON, defining recommender systems to support software development and design of hybrid persistence systems.

References:

[1]. Phuong T. Nguyen; Claudio Di Sipio; Juri Di Rocco; Massimiliano Di Penta; Davide Di Ruscio, "***Adversarial Attacks to API Recommender Systems: Time to Wake Up and Smell the Coffee?***" in Proceedings of the 36th IEEE/ACM International Conference on Automated Software Engineering (ASE), DOI: 10.1109/ASE51524.2021.9678946

[2]. Phuong T. Nguyen, Davide Di Ruscio, Juri Di Rocco, Claudio Di Sipio, Massimiliano Di Penta, "***Adversarial Machine Learning: On the Resilience of Third-party Library Recommender Systems***" in Proceedings of the International Conference on Evaluation and Assessment in Software Engineering (EASE 2021), DOI: 10.1145/3463274.3463809

==========

TS. Nguyễn Thanh Phương tốt nghiệp Kỹ sư và Thạc sĩ CNTT, Đại học Bách Khoa Hà Nội năm 2002 và 2004. Sau đó năm 2012, anh Phương lấy bằng Tiến sĩ Khoa học Máy tính tại Đại học Tổng hợp Jena, CHLB Đức. Khi về nước, anh tham gia giảng dạy hệ đại học, cao học cho Đại học FPT, Đại học Duy Tân. Trong khoảng thời gian một năm (2014 đến 2015), TS. Phương làm nghiên cứu sinh sau tiến sĩ tại Đại học Bách Khoa Bari, Cộng hòa Ý. Từ năm 2017, TS. Phương tiếp tục làm nghiên cứu sinh sau tiến sĩ tại Đại học Tổng hợp L'Aquila, Cộng hòa Ý, và cho tới năm 2022 thì được nhận vị trí giáo sư tập sự (Tenure track assistant professor) cũng tại đó.

Các chủ đề nghiên cứu mà TS. Phương thực hiện là về hệ khuyến nghị (Recommender Systems), Học máy và học sâu (Machine Learning, Deep Learning), và ứng các dụng trong ngành Công nghệ phần mềm (Software Engineering) đặc biệt là để khai thác các kho dữ liệu mã nguồn mở (GitHub, Maven, Stack Overflow), Chẩn đoán hình ảnh Y khoa (Diagnostic Imaging). Anh Phương đã xuất bản hơn 50 bài báo, trong đó có nhiều bài trên các tập san và hội thảo quốc tế được xếp hạng cao như IEEE Transactions on Software Engineering (TSE), Elsevier Journal of Systems and Software (JSS), Elsevier Expert Systems with Applications (ESWA), International Conference of Software Engineering (ICSE), và International Conference on Automated Software Engineering (ASE).

Ngoài công việc nghiên cứu, TS. Phương là Phó Tổng biên tập (Associate Editor) của tập san Applied Intelligence (Scimago Q2, IF=4.760) (https://www.springer.com/journal/10489/editors), và là thành viên của Ban Biên tập The Journal of Universal Computer Science (Scimago Q2) (https://bit.ly/3NRnsve). TS. Phương tham gia giảng dạy nhiều môn, trong đó có môn Phương pháp nghiên cứu khoa học về cách thức viết bài báo, trả lời bình duyệt.

Các chủ đề chính sẽ trình bày trong buổi seminar tại VIASM, ngày 18/08/2022.

1. Một số kỹ thuật thay đổi dữ liệu để làm sai lệch (gây tác động xấu) tới hệ khuyến nghị trong Công nghệ phần mềm (Recommender Systems in Software Engineering).
2. Một số chủ đề nghiên cứu đang được quan tâm hiện nay trong Công nghệ phần mềm.
3. Thông tin về một số chương trình học bổng tiến sĩ và sau tiến sĩ tại Ý.