Paying Attention to the Insider Threat

Eduardo Lopez
Information Systems, DeGroote School of Business
McMaster University
Hamilton, ON, Canada
lopez1@mcmaster.ca

Kamran Sartipi
Department of Computer Science
East Carolina University
Greenville, NC, USA
sartipik16@ecu.edu

Abstract—The misuse of information systems by internal actors – the insider threat – is an ever-growing concern in organizations of all types. The timely detection of an insider threat is as important as it is difficult. Analyzing user behaviours recorded in electronic logs require significant computing resources and the capability to find and interpret complex patterns in temporal sequences that may contain irrelevant, temporary or novel elements. In this paper we use an attention-based architecture derived from BERT (Bidirectional Encoder Representations from Transformers) for the creation, storage and updating of an always-current, holistic user behaviour model that enables real-time insider threat detection through anomaly detection and user behaviour prediction techniques.

Keywords: insider threat, transformers, BERT, anomaly detection, cybersecurity

I. INTRODUCTION

Information technologies enable human endeavours like never before. In developed societies the use of information systems is pervasive – from individuals that use smartphones to organizations that manage the planning and allocation of resources. The significant benefits from advanced information technology also come with an unintended byproduct which is the potential for information system misuse by authorized users or impersonated people. We refer to this phenomena as the insider threat. A current statistic reveals the magnitude of the problem as 34% of data breaches in 2018 involving internal actors [20].

Detecting an on-going insider threat is a significant challenge. Although the actions by every user are regularly recorded in electronic files (i.e., logs), those logs can be obscured by a very large number of unrelated events. Information in electronic logs is usually unstructured, stored in very large text files that require specialized tools to be analyzed. User behaviours are captured in sequences of events that can be mined for abnormal patterns. However, their ever-evolving characteristic and dependence on myriad contextual parameters (e.g., the time of use or location of computer) pose remarkable challenges for advanced analytic applications.

Deep Learning – neural networks with multiple hidden layers – are powerful function approximators that can be used in insider threat detection [10]. Over the preceding decade, Recurrent Neural Networks (RNN) of many types have been utilized for this purpose, with some architectures – such as Long Short-Term Memory (LSTM) – achieving remarkable results in the analysis of event sequences. For the purpose of this work, we focus on a set of technologies known as Transformers [19] that have dramatically improved Artificial Intelligence (AI) applications in Natural Language Processing (NLP). They are encoder/decoder architectures that implement the concept of Attention.

The preceding decade has witnessed remarkable advances in Artificial Intelligence (AI). Deep learning in particular has consistently delivered results across myriad disciplines, sometimes surpassing human-level benchmarks [1]. A deep learning architecture is a multi-layer stack of neural network units where usually most of them are subject to learning and that may include non-linear input-to-output mappings [10]. Each layer in the network incrementally learns about the structure of data from its preceding layers, becoming very good function approximators that can find and learn very complex patterns in the data. There are several deep learning architectures that can be applied to specific applications and data types.

Deep learning has been proven effective across a wide range of disciplines, with cybersecurity being the focus of this work. The protection of information assets against malicious threats is a pivotal element for an increasingly technified society where information systems enable processes for many organizations. Arguably, one of the most interesting challenges pertains to the phenomena known as the insider threat. It can be defined as current (or former) users – or somebody impersonating them – that intentionally misuse access privileges negatively impacting the confidentiality, integrity or availability of information or information systems [6].

In this study we train a user behaviour model as the baseline to perform anomaly detection and behaviour prediction for insider threat detection. More specifically, the main contributions of this work are as follows:

- A Transformers-based architecture that uses self-attention for modeling user behaviours. The model created can be used effectively in transfer learning tasks.
- Design of the data engineering and Data Science / Machine Learning (DSML) pipelines to enable timely and consistent insider threat detection.
- Two complementary techniques for identifying information systems misuse, including: time-series anomaly detection using features extracted from the DSML pipeline, and user behaviour prediction.