Detecting Insider Threat from Large Transaction Datasets using LSTM

Eduardo Lopez
DeGroote School of Business
McMaster University
Hamilton, ON, Canada
lopeze1@mcmaster.ca

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

***Abstract*—Information technologies enable business processes in most organizations. Notwithstanding its benefits, the technification of society's processes come with an unintended byproduct: the opportunity for attackers to misuse information systems, sometimes impersonating authorized users – a phenomena referred to as the insider's threat. Detecting an insider's threat taking place may be achieved by analyzing logs containing system's events and user behaviours. However, logs are usually very large and unstructured files, difficult to read and resource-intensive to analyze. In this paper, we use deep learning and most specifically Long Short Term Memory (LSTM) recurrent networks for enabling the detection of insider attacks. We demonstrate how the use of an bidirectional LSTM architecture is capable of modelling user behaviour and enable the cost-effective detection of the threat.**

Keywords: insider threat, LSTM, cybersecurity, feature engineering, time-based sequences.

I. Introduction

A significant amount of resources – financial and human – are invested every year in cybersecurity. The 2019 expenditure exceeded $120 billion [2], but the growth of attacks continues unabated [4]. Many of these attacks involve the use of existing accounts, either by existing users or somebody impersonating them. We refer to this phenomena as the insider's threat.

Detecting an insider's threat taking place is a difficult task. Although electronic logs record the threat taking place, their size and complexity pose significant challenges for a timely analysis. User behaviour – the sequence of events that a user performs in an information system – may be mined for irregular patterns that may indicate the treat taking place. A user behaviour that diverges from a prediction may provide useful information when detecting a threat. The prediction shall account for the characteristics of user behaviour: dynamic, changing over time or depending on myriad contextual factors such as the physical location of the user, the day in the week or even the device used.

For the purposes of this work, we focus on a set of technologies that have yielded remarkable results: deep neural networks, commonly referred to as Deep Learning. It pertains to the use of neural networks with multiple hidden layers that have achieved performance levels beyond human benchmarks across multiple tasks [19]. Deep learning discovers data patterns and structures by training with large data sets changing its internal parameters through a back-propagation algorithm [13]. Some deep learning architectures exploit the sequential nature of the data for finding the patterns through specialized networks known as Recurrent Neural Networks or RNNs. Our approach uses a type of RNN well-suited for analysis of long sequences: Long Short-Term Memory (LSTM). We explore how the use of certain LSTM architectures is best suited for the successful detection.

The main contributions of this work are as follows:

- We articulate a structured and innovative approach to encoding, unsupervised training, predicting and evaluating a deep learning model for insider's threat detection that is based on user behaviour prediction based on both preceding and succeeding behaviours.
- We demonstrate the approach with a very large multi-system, multi-user log that includes ground truth records capturing red team activities. This is an important contribution given the scarcity of datasets available for insider threat testing [25].
- We identify the critical factors that practitioners should consider when designing and implementing deep learning-based systems for insider threat detection.

This paper is organized as follows: in Section II we document relevant research previously performed. In Section III we explain the theoretical framework for understanding the foundations of neural networks, deep learning and recurrent neural networks. Section IV describes how we architect anomaly detection through the selected hardware and software, followed by the experiments in Section V. We close the discussion with our conclusions and suggested research streams in Section VI.

II. Related Work

Detecting the insider threat in information systems is a widely researched area with different effective approaches. A useful categorization to understand extant research is whether the approach uses labelled data (i.e., supervised learning) or assumes no knowledge of a signature to differentiate regular information systems use from insider threats. For the purposes of this study, the focus is unsupervised learning so there is alignment with conditions found in practice. A second category that pertains to this work is the use of deep learning for the detection. Notwithstanding the paucity of studies in