Nom et Prénom : BOUKHALFA ALAEDDINE
Date de soutenance : 25/12/2020
Directeur de Thèse : H. EL FADIL

Sujet de Thèse :

Détection des intrusions à l'aide de la méthode du Deep Learning LSTM dans un environnement Big Data

Résumé :

De nos jours, les anciennes villes se sont transformées en ville intelligentes qui manipulent les nouvelles technologies de l'information et de la communication (NTIC) afin d'améliorer les performances des services urbains et réduire les coûts.

Les dispositifs électroniques utilisés au sein de ces villes intelligentes communiquent entre eux via un nouveau réseau mondial appelé internet des objets (IdO) et sont à l'origine de la production d'une grande masse et une variété des données.

En plus, les communications entre ces dispositifs électroniques ne sont pas toujours sécurisées, et ils peuvent cacher d'une manière discrète de nouvelles intrusions de différents types.

Or, les outils de la sécurité informatique qui existants, sont établis sur des algorithmes non évolutifs pour détecter les nouvelles intrusions, et ont du mal à les détecter dans un environnement de grande quantité et variété des données.

Dans cette thèse, nous proposons, de nouvelles approches basées sur les techniques du Big Data et les algorithmes du Machine Learning et Deep Learning (DL) afin de résoudre les problèmes cités ci-dessus, et améliorer la sécurité de l'information dans un environnement à grande masse et variété des données.

La première approche consiste à mettre en place une architecture de monitoring de la sécurité du système d'information, qui s'évolue par apprentissage pour identifier les nouvelles intrusions de différents types, dans un environnement où les données sont volumineuses et diversifiées. Le principe de l'architecture est basé sur la collecte de la grande quantité et la variété des données des intrusions, leur stockage en utilisant les techniques du Big Data, et leur analyse par la méthode du Deep Learning Recurrent Neural Network (RNN), afin de reconnaître ces intrusions, pour pouvoir bloquer d'autres nouvelles intrusions qui peuvent apparaitre. L'approche a été évaluée. Les résultats ont confirmé que l'architecture proposée est très efficace et évolutive pour la détection de la variété des nouvelles intrusions. Au moment des expérimentations de cette première contribution, nous avons remarqué que l'analyse des données des intrusions consomme un temps considérable, ce qui a nécessité une amélioration.

La deuxième approche repose sur l'amélioration de la première approche tout en proposant une architecture de traitement en parallèle et distribué via un cluster Big Data avec plusieurs nœuds, dans le but de réduire le temps nécessaire pour l'analyse des données des intrusions. L'expérimentation a été effectuée. Les résultats ont montré que le traitement en parallèle et distribué réduit considérablement la consommation du temps lors de l'analyse des données des intrusions..

Mots-clés :
Monitoring de la Sécurité, Détection des Intrusions, Machine Learning, Deep Learning, SVM, Arbre de Décision, KNN, CNN, LSTM, RNN, Big Data.

Abstract :

Nowadays, old cities was transformed into smart cities that manipulate new information and communication technologies (NICT) to improve the performance of urban services, reduce costs and minimize resource consumption.

The electronic and digital devices used within these smart cities communicate with each other through a new global network called the Internet of Things (IoT) and are responsible for the production of a large mass and variety of data.

In addition, communications between these electronic devices are not always secure, and they can discreetly hide new attacks and intrusions of different types.

However, computer security mechanisms and tools that currently exist, are based mainly on algorithms and methods that are not evolutionary to detect new attacks and intrusions, and also suffer with the large quantity and variety of data produced.

In this thesis, we propose new approaches based on Big Data techniques and Machine Learning and Deep Learning algorithms in order to solve the problems mentioned above, and improve information security within an environment of large mass and variety of data.

The first approach consists in setting up an information system security monitoring architecture, that evolves by learning to identify new intrusions of different types, in an environment where the data is large and diverse. The principle of the architecture is based on the collection of the large quantity and variety of intrusion data, their storage using Big Data techniques, and their analysis by the Recurrent Neural Network (RNN) Deep Learning method, to recognize these intrusions, in order to block other new intrusions. The approach was evaluated. The results confirmed that the proposed architecture is very efficient and evolutionary for the detection of the variety of new intrusions. At the time of the experiments of this first contribution, we noticed that the analysis of intrusion data consumes a considerable amount of time, which required improvement.

The second approach is based on improving the first approach proposing a parallel and distributed processing architecture through a Big Data cluster with multiple nodes, in order to reduce the time required for intrusions data analysis. The experiment was carried out. The results showed that parallel and distributed processing reduces considerably time consumption during intrusion data analysis.

Keywords :
Security Monitoring, Intrusion Detection, Machine Learning, Deep Learning, SVM, Decision Tree, KNN, LSTM, RNN, CNN, Big Data.