Offre n°2023-07065

Post-Doctoral Research Visit F/M Data Augmentation for Reliable training of ML-based Intrusion Detectors

Le descriptif de l’offre ci-dessous est en Anglais

Type de contrat : CDD

Niveau de diplôme exigé : Thèse ou équivalent

Fonction : Post-Doctorant

Contexte et atouts du poste

The offered position is proposed by the RESIST team of the Inria Nancy Grand Est research lab, the French national public institute dedicated to research in digital Science and technology. The team is one of the European research group in network management and is particularly focused on empowering scalability and security of networked systems through a strong coupling between monitoring, analytics and network orchestration.

https://team.inria.fr/resist/

This work is in the context of the SuperviZ project. The SuperviZ project is part of the "system security" axis of the PEPR cybersecurity program. It addresses the field of "system, software and network security". More precisely, it targets the detection, response and remediation to computer attacks, subjects grouped under the name of "security supervision".

Mission confiée

Cybersecurity is a major concern everywhere with the growth of connected devices that are beyond common computers. To circumvent these problems, decades of research and development have led to build new techniques and tools to fight back against the attacks on the Internet. Nonetheless, the number of attacks and their magnitude still grow. The attack surface continues to increase along with the number of connected devices but also due to the number of applications, services or software that today make the IT ecosystem far from its origin.

Techniques used by both attackers and defenders evolve to complex mechanisms. For example, this leads to the massive use of encryption to avoid data leaks but simultaneously attackers benefit from encryption to hide their own activities. Multiple steps attacks also requires to analyze numerous sources of data. As a result, intrusion detection methods relying on artificial intelligence have been investigated both in research and in industry.

While these techniques hold promise for detecting and mitigating cyber threats, their effectiveness is highly dependent on the quality of the learning phase. Despite significant progress, experiments and reports suggest that these tools still struggle to generalize effectively to new and previously unseen data, particularly when faced with minor variations compared to training data.

Actually, the learning suffers from the lack of enough labeled data to represent the different and possibly infinite variation of attacks. To avoid this problem, different approaches exist. Among them, data augmentation consists into extending artificially the set of input data for learning in a realistic way.

Principales activités

The aim of the postdoc is to assess the effectiveness of data augmentation techniques in enhancing the robustness of attack detection mechanisms based on machine learning classifiers. Concretely, it consists in extending datasets of network traffic containing attacks and evaluate the accuracy of the ML classifier with the newly generated data (with or without retraining).

The main challenge is to identify an appropriate data augmentation technique that is relevant to our context. Although Generative Adversarial Networks (GANs) [1] have been widely used to produce models by automating the generation of data, they are susceptible to result in artificially generated data with limited variation and reduced value. Therefore, different improvements have been proposed [2, 3]. In our context, applying a common GAN architecture has been proved to be inefficient and authors in [4] propose to decompose input data into multiple groups before applying a GAN, in order to augment data for each individual type of attacks.

In addition to GANs, alternative approaches such as using a well-selected sequence of data transformations, also known as a data augmentation policy, have been explored [5].
The baseline technique promoted in [4] will be used as a reference, and the relevant possible data transformations of network traffic traces and strategies will be defined to compare this technique with a transformation-based augmentation technique. The postdoctoral researcher would have to identify relevant transformation for network traces considering the different layers including application payload. A formal definition of these transformations will be leveraged to define an augmentation policy engine. To determine the most appropriate data augmentation policy, an iterative strategy based on reinforcement learning will be proposed.

In addition, the postdoc will have the opportunity to experiment with other approaches, such as using adversarial autoencoders or Kronecker Graphs.

References

Compétences
- Required qualification: PhD in Computer Science
- Required knowledge: networking, network security, machine learning including practical experiences with large datasets
- Languages: Shell, python, ML libraries and others are appreciated
- Fluent in english (writing and oral communication)

Avantages
- Subsidized meals
- Partial reimbursement of public transport costs
- Leave: 7 weeks of annual leave + 10 extra days off due to RTT (statutory reduction in working hours) + possibility of exceptional leave (sick children, moving home, etc.)
- Possibility of teleworking (after 6 months of employment) and flexible organization of working hours
- Professional equipment available (videoconferencing, loan of computer equipment, etc.)
- Social, cultural and sports events and activities
- Access to vocational training
- Social security coverage

Rémunération
2788€ gross/month

Informations générales
- Thème/Domaine : Réseaux et télécommunications
- Système & réseaux (BAP E)
- Ville : Villers lès Nancy
- Centre Inria : Centre Inria de l'Université de Lorraine
- Date de prise de fonction souhaitée : 2024-03-01
- Durée de contrat : 2 ans
- Date limite pour postuler : 2024-01-21

Contacts
- Équipe Inria : RESIST
- Recruteur : François Jerome / jerome.francois@inria.fr

A propos d'Inria
Inria est l'institut national de recherche dédié aux sciences et technologies du numérique. Il emploie 2600 personnes. Ses 215 équipes-projets agiles, en général communes avec des partenaires...
académiques, impliquent plus de 3900 scientifiques pour relever les défis du numérique, souvent à l'interface d'autres disciplines. L'institut fait appel à de nombreux talents dans plus d'une quarantaine de métiers différents. 900 personnels d'appui à la recherche et à l'innovation contribuent à faire émerger et grandir des projets scientifiques ou entrepreneuriaux qui impactent le monde. Inria travaille avec de nombreuses entreprises et a accompagné la création de plus de 200 start-up. L'institut s'efforce ainsi de répondre aux enjeux de la transformation numérique de la science, de la société et de l'économie.

**L'essentiel pour réussir**

Applications are to be sent as soon as possible.

Tentative starting date: October 2023

How to apply:

Upload your file on jobs.inria.fr in a single pdf or zip file, and send it as well by email to jerome.francois@inria.fr and isabelle.chrisment@inria.fr. Your file should contain the following documents:

- Your CV.
- A cover/motivation letter describing your interest

In addition, one recommendation letter from the person who supervises(d) your work should be sent directly by his/her author to jerome.francois@inria.fr and isabelle.chrisment@inria.fr.

**Attention**: Les candidatures doivent être déposées en ligne sur le site Inria. Le traitement des candidatures adressées par d'autres canaux n'est pas garanti.

**Consignes pour postuler**

**Sécurité défense**:

Ce poste est susceptible d'être affecté dans une zone à régime restrictif (ZRR), telle que définie dans le décret n°2011-1425 relatif à la protection du potentiel scientifique et technique de la nation (PPST). L'autorisation d'accès à une zone est délivrée par le chef d'établissement, après avis ministériel favorable, tel que défini dans l'arrêté du 03 juillet 2012, relatif à la PPST. Un avis ministériel défavorable pour un poste affecté dans une ZRR aurait pour conséquence l'annulation du recrutement.

**Politique de recrutement**:

Dans le cadre de sa politique diversité, tous les postes Inria sont accessibles aux personnes en situation de handicap.