sTressINg sysTems sEcurity thRough on the Fly nEtwork tRaffic gEneration

Stage sous la direction de Francesco Bronzino (maître de conférence à l'ENS Lyon, laboratoire LIP, équipe HoWNet) et Antoine Boutet (maître de conférences de l'INSA Lyon, laboratoire CITI, équipe Inria Privatics)

Thèmes : Intelligence Artificielle / Données synthétiques / Sécurité Informatique

Mots clés : IA, Données synthétiques, Vie privée

Contexte

Les données séquentielles - données qui dépendent du temps - sont très courantes, allant des transactions par carte de crédit aux dossiers médicaux en passant par les cours boursiers. Mais les réglementations en matière de confidentialité d'informations personnelles limitent et ralentissent considérablement l'accès aux données utiles, essentielles à la recherche et au développement. Cela crée une demande de génération de données séquentielles synthétiques à la fois hautement représentatives et préservant la confidentialité.

Les données synthétiques permettent également de répondre à certaines spécificités. La rareté de certains événements, ou leur faible importance relative peut rendre leur anticipation difficile, alors que la modularité des données synthétiques permet de générer toutes sortes de situations. Il est possible de générer un événement difficile à enregistrer en pratique, ou relativement rare, comme une fraude fiscale, une attaque sur un système d'information, ou une maladie rare. Cette flexibilité rend les données synthétiques extrêmement utiles pour créer des outils d'audit évaluant la réponse de systèmes à des évènements spécifiques.

Problème de recherche

Dans le projet INTERFERE, on se concentre sur la génération de données séquentielles de type trafic réseau ou flux de requêtes permettant de stresser des systèmes face à des problèmes de sécurité. On considère notamment les tentatives d'intrusion ou des tentatives de ré-identification d'utilisateurs à travers de multiples requêtes sur un entrepôt de données anonymes. Le projet INTERFERE vise à générer de grosses quantités de données à la volée avec un contrôle des événements représentés. En proposant une recherche d'outils d'audit, ce projet joue un rôle clé dans le développement de nouvelles méthodes et technologies pour renforcer la sécurité des systèmes et des données.

Etat de l'art

La génération de séries chronologiques synthétiques et de données séquentielles est plus difficile que les données tabulaires où normalement toutes les informations concernant un individu sont stockées sur une seule ligne. Dans les données séquentielles, les informations peuvent être réparties sur plusieurs lignes, comme les transactions par carte de crédit, et la préservation des corrélations entre les lignes - les événements - et les colonnes - les variables - est essentielle. De plus, la longueur des séquences est variable ; certains cas peuvent ne comprendre que quelques transactions tandis que d'autres peuvent en avoir des milliers.

Bien que les modèles génératifs pour les données séquentielles et les séries chronologiques aient été étudiés, bon nombre de ces efforts ont abouti à des résultats dont l'applicabilité est encore limitée dans la pratique [3]. De plus, cette qualité rentre en conflit dans certains cas avec le niveau de confidentialité des données générées qui peuvent faire fuiter des informations personnelles liées aux données d'origine [2,4].

Les travaux existants sur les générateurs de trafic peuvent être globalement classés en deux catégories : (1) les générateurs de traces de trafic synthétiques [2-4] et (2) les générateurs de trafic en temps réel. La première catégorie a pour objectif principal la génération hors ligne de nouvelles traces à partir d'un jeu de données de traces existant. Des travaux récents [3, 4] ont adopté les GAN comme une solution efficace pour générer des traces qui reproduisent fidèlement les caractéristiques de la trace d'origine tout en assurant la confidentialité (par exemple, en changeant les adresses IP) de l'ensemble de données nouvellement créé. La deuxième catégorie a pour objectif de concevoir des systèmes capables d'injecter du trafic en temps réel dans un réseau. Comme le trafic injecté doit ressembler à un trafic réaliste, ces travaux mettent en oeuvre différentes stratégies pour piloter la manière dont le trafic est généré au moment de l'exécution. Ces stratégies vont de l'utilisation de solutions basées sur des scripts [5], la relecture de traces existantes [6], ou des solutions hybrides [7].

Contrairement aux travaux antérieurs, notre ambition est de tirer les leçons des deux approches, en visant une solution qui puisse à la fois créer un trafic réaliste - c'est-à-dire en utilisant des modèles ML complexes - et en temps réel - c'est-à-dire de la même manière que les générateurs de trafic. De plus, on souhaite également tester les systèmes dans une variété de scénarios en contrôlant les événements à mettre en avant à des fins d'audit.

Objectifs: L'objectif du stage vise à développer des outils d'audit exploitant la génération de données synthétiques. Les principaux objectifs du projet sont les suivants :

- Développer des méthodes d'apprentissage de modèles génératifs offrant une qualité de données permettant de générer du trafic réseau ou des flux de requêtes hautement représentatif.
- Développer des méthodes de génération de données synthétiques permettant de générer du trafic réseau ou des flux de requêtes en temps réel et en grosse quantité. Nous nous concentrerons sur le passage à l'échelle de cette génération et comment envoyer les données générées sur le réseau.
- Développer des méthodes permettant de faire varier à la demande les données générées afin de faire apparaître des événements particuliers tels que des tentatives d'intrusion ou de ré-identification d'utilisateurs.

Prérequis et compétences attendues :

- Compétence en programmation
- Connaissance en apprentissage machine
- Bonne aptitude à collaborer et à communiquer (écrit / oral)
- Autonome, force de proposition
- Un intérêt pour la sécurité / la protection des données personnelles serait un plus

Structure d'accueil : Ce stage sera hébergé par l'équipe HoWNet, dans les locaux de l'ENS sur le campus de Gerland. Ces travaux seront réalisés en collaboration avec Francesco Bronzino et Antoine Boutet. Contacts : francesco.bronzino@ens-lyon.fr, antoine.boutet@insa-lyon.fr

Références

[1] **Federated Learning: Strategies for Improving Communication Efficiency.** Jakub Konečný, H. Brendan McMahan, Felix X. Yu, Peter Richtárik, Ananda Theertha Suresh, Dave Bacon. ArXiv, cs.LG, 1610.05492, 2016.

[2] **Salvaging Federated Learning by Local Adaptation.** Tao Yu, Eugene Bagdasaryan, Vitaly Shmatikov. ArXiv, cs.LG, 2002.04758, 2020.

[3] **Beyond Inferring Class Representatives: User-Level Privacy Leakage From Federated Learning.** Zhibo Wang, Mengkai Song, Zhifei Zhang, Yang Song, Qian Wang, Hairong Qi. INFOCOM, 2019.

[4] **How To Backdoor Federated Learning.** Eugene Bagdasaryan, Andreas Veit, Yiqing Hua, Deborah Estrin, Vitaly Shmatikov. ArXiv, cs.CR, 1807.00459, 2018.

[5] **Inferring Sensitive Attributes from Model Explanations.** Vasisht Duddu, Antoine Boutet. CIKM 2022.

[6] **MIXNN: Protection of Federated Learning Against Inference Attacks by Mixing Neural Network Layers.** Thomas Lebrun, Jan Aalmoes, Adrien Baud, Antoine Boutet. Middleware 2022.

[7] **Privacy Assessment of Federated Learning using Private Personalized Layers.** Theo Jourdan, Antoine Boutet, Carole Frinder. MLSP 2021.