Titre : Application de techniques d'apprentissage pour la gestion des accès

Encadrant : Clara Bertolissi (clara.bertolissi@univ-amu.fr), équipe MoVe

Mot clés : gestion des accès, machine learning, vérification des politiques, détection d'anomalies

Gérer les accès aux données sensibles est un enjeu primordiale dans les systèmes numériques actuels. Les droits d'accès sont généralement exprimés sous forme des règles qui constituent ce qu'on appelle une politique de contrôle d'accès.
L'étape de vérification d'une politique de contrôle d'accès garantit qu'il n'y a aucune erreur dans la politique, autrement dit, que le module de gestion des accès se comporte comme attendu.
La vérification de politiques de contrôle d'accès s'appuie traditionnellement sur des méthodes telles que la preuve de modèles, la simulation de systèmes ou le test pour vérifier que la logique sous-jacente à la politique fonctionne de manière cohérente.
Cependant, ces méthodes présentent des problèmes de capacité et de performances liés aux technologies appliquées.

Récemment, plusieurs travaux de recherche ont proposé des solutions pour le contrôle des accès basé sur des techniques d'apprentissage (ML) [Survey].
Par exemple, plusieurs travaux proposent des solutions basées sur le ML pour concevoir des politiques de contrôle d'accès plus robustes que les approches classiques.
Ces dernières années ont également vu l'utilisation du ML pour la vérification des politiques pour en améliorer la qualité lors de leur mise en œuvre. L'arbre de décision (DT) et la classification aléatoire des forêts (RFC) sont deux des principaux algorithmes de classification ML capables de vérifier efficacement les politique de contrôle d'accès.
En particulier, dans le rapport du NIST de Vincent C. Hu [NIST], le modèle de ML utilisé est entrainé
avec des données qui encodent les règles de la politique de contrôle d'accès dans une table de données. Le modèle est ensuite utilisé pour prédire les autorisations d'accès (accorder ou refuser certains droits) en indiquant les défauts trouvés dans les règles de la politique et en permettant ainsi la détection d'incohérences.
Cette méthode de vérification ne nécessite pas de cas de test complets, ou de traduction du système en un modèle formel, mais vérifie plutôt la logique des

règles de la politique directement, ce qui le rend plus efficaces et réalisables par rapport aux méthodes traditionnelles.

Le bût de ce stage est de se familiariser avec ce type d'application du ML à la gestion des accès, en particulier en se basant sur [NIST] et en utilisant l'algorithme RFC [SKLEARN] sur un example de politique à vérifier.

Ce stage peut se poursuivre en thèse dans le cadre du Projet TrustinClouds, où les modèles et mécanismes de gestion des accès sont appliquée aux environnement cloud. Des techniques de ML seront investiguées pour pour apprendre les profils comportementaux des utilisateurs accédant aux resources, dans le bût d'affiner dynamiquement les politiques de contrôle d'accès et pour détecter des abus de privilèges de la part des utilisateurs.

[Survey] Machine Learning in Access Control: A Taxonomy and Survey. 2022. Mohammad Nur Nobi, Maanak Gupta, Lopamudra Praharaj, Mahmoud Abdelsalam, Ram Krishnan, Ravi Sandhu. https://arxiv.org/abs/2207.01739

[NIST] Vincent Hu. 2021. Machine Learning for Access Control Policy Verification. Technical Report. https://doi.org/10.6028/NIST.IR.8360

[SKLEARN] scikit-learn (2020) sklearn.ensemble.RandomForestClassifier, scikit-learn 0.24.1. Available at https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestClassifier.html