Un nouveau système immunitaire artificiel pour l'apprentissage non supervisé

Rachid Elmeziane (*), Ilham Berrada (*) et Ismail Kassou (*)

* Laboratoire Al Khawarizmi- ENSIAS- BP 713- Agdal –Rabat- Maroc
(*){meziane, iberrada, kassou}@ensias.ma

Résumé. Nous proposons dans ce papier un nouveau système immunitaire artificiel (SIA) appelé système NK, pour la détection de comportement du soi non soi avec une approche non supervisée basée sur le mécanisme de cellule NK (Naturel Killer). Dans ce papier, le système NK est appliqué à la détection de fraude en téléphonie mobile.

1 Contexte

Dans le but de résoudre des problèmes complexes du monde réel dans des domaines différents tels que l'optimisation, la détection d'anomalies ou la robotique, des heuristiques inspirées de mécanismes naturels ont été exploitées avec succès. Plusieurs chercheurs se sont intéressés aux systèmes immunitaires biologiques (SIB) comme un nouveau paradigme de l'intelligence artificielle et ont développé des applications industrielles en ordonnancement, en robotique, ou en détection d'intrusion. Néanmoins, peu de travaux ont traité la problématique de la détection de fraude de comportement en télécommunications.

Dans ce papier, on propose un nouveau système immunitaire artificiel (SIA) pour la détection du comportement du soi non soi avec une approche non supervisée basée sur le mécanisme SIB dit inné de cellule NK. Un tel système diffère des SIA existants qui se basent sur le mécanisme supervisé adaptatif de SIB des cellules T et B (Garrett 2005).

2 Présentation du système NK proposé

L'algorithme de notre système NK, décrit dans le tableau TAB1, comporte quatre phases qui concernent la reconnaissance et l'extraction de modèles d'instances puis leur transformation en signaux d'inhibition et d'activation. La dernière phase concerne la détection de la présence de comportements anormaux sur la base de l'analyse des densités spectrales ou de filtrage des signaux. Notons ici que la terminologie signal utilisée correspond à un signal discret à temps discret et que l'entrée de l'algorithme est une série chronologique vectorielle.

L'algorithme du système NK élaboré a été testé sur des données simulées de 10100 instances de télécommunication, relatives aux trafics de certains usagers chez un intermédiaire, et qui sont infectées par un comportement frauduleux pour les instances entre 10000 et 10100 et dont la proportion représente 0.01% de l'échantillon. Les résultats obtenus sont satisfaisants car, malgré la proportion très faible des opérations frauduleuses dans l'échantillon, notre système NK a réussi à les détecter (cf. FIG. 3) et à identifier les instances de comportements frauduleux (cf. FIG. 1 et FIG. 2). Ces mêmes résultats seront comparés avec

d'autres algorithmes de classification non supervisés tels que les réseaux de Kohonen, K-Means et autres.

Phase 1 : **Reconnaissance et extraction de modèles d'instances**

- Représentation des données de comportement en forme binaire.
- Représentation des modèles d'instances de comportement par des opérations de permutations de l'algorithme GRP utilisé en cryptographie (Lee, Shi et Yang 2001).

Phase 2 : **Modélisation du signal d'inhibition** sig_kir

- Génération du signal d'inhibition par l'entropie de permutation (Bandt et Prompe 2002).

Phase 3 : **Modélisation du signal d'activation** sig_kar

- Génération du signal d'activation par la distance Damerau-Levenshtein (Hyyrö 2003)

Phase 4 : **Détection de comportements anormaux**

- Analyse de chaque signal d'inhibition sig_kir et d'activation sig_kar par l'algorithme MUSIC pour produire leur densité spectrale PSD.
- Détection de la présence d'une zone de comportements anormaux par filtrage et par différenciation de PSD de chaque signal.

TAB. 1 - *Les différentes phases du système NK proposé*

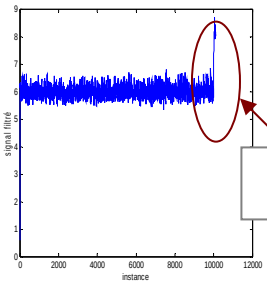

FIG. 1 - *sig_kir filtré*

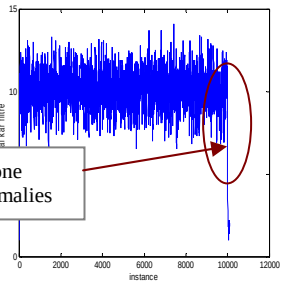

FIG. 2 - *sig_kar filtré*

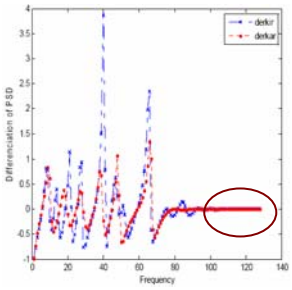

FIG. 3 - *Dérivées des PSD du sig_kir et sig_kar*

Références

Bandt, C. et B. Prompe (2002). Permutation entropy - a natural complexity measure for time series. *Phys. Rev, Lett. 88*, 174102.

Hyyrö, H. (2003). A bit-vector algorithm for computing levenshtein and damerau edit distances *Nord. J. Comput. 10(1)*, 29–39.

Lee, R. B., Z. Shi, et X.Yang (2001). Efficient permutation instructions for fast software cryptography. *IEEE Micro 21(6)*, 56–69.

Summary

We propose a new artificial immune system called NK system for the detection of self non self behaviour with an unsupervised approach based on the mechanism of NK cell. In this paper, the NK system is applied to the detection of fraud in mobile telephony.