heig-vd
Haute Ecole d'Ingénierie et de Gestion
du Canton de Vaud

Tél. : 024 557 63 30
Fax : 024 557 64 04

info@heig-vd.ch
http://www.heig-vd.ch

Travail de diplôme 2005

Département Electricité et Informatique

HEIG-VD
Route de Cheseaux 1
CH - 1401 Yverdon-les-Bains

Système d'aide lors d'enquêtes sur des cas d'escroquerie par e-mail

Reconnaissance automatique de la structure d'e-mails

Le phénomène de la « Nigerian Connection » (nommé également 419 scam) est un type d'escroquerie par e-mail, qui est actif depuis plusieurs années, et qui se base désormais sur les technologies de l'Internet.

Le mode opératoire des escrocs consiste à proposer par exemple une forte récompense ou une grosse somme d'argent en échanges de services, ceci par l'envoi massif d'e-mails dans le but d'appâter des internautes. L'une des caractéristiques de cette fraude est l'évolution constante des scénarios. Une fois le contact établi avec une victime potentielle qui a eu l'imprudence de répondre au courriel, le malfaiteur en vient peu à peu à demander des avances de frais sous divers prétextes (frais de procédure, frais de transfert, etc.).

Les citoyens qui flairent l'arnaque transmettent parfois ces messages (ou spams) à la police. Ces derniers constituent une précieuse source d'informations, souvent inexploitées.

Ce projet de diplôme a pour but de développer un système automatisé d'acquisition, de structuration, de sauvegarde et d'analyse de ces e-mails (figure 1). Il se base sur deux rapports successifs de l'Institut de Police Scientifique traitant de ce phénomène.

L'architecture du système à développer se veut souple et modulaire afin d'intégrer les évolutions rapides des scénarios et les nouveaux types de fraudes.

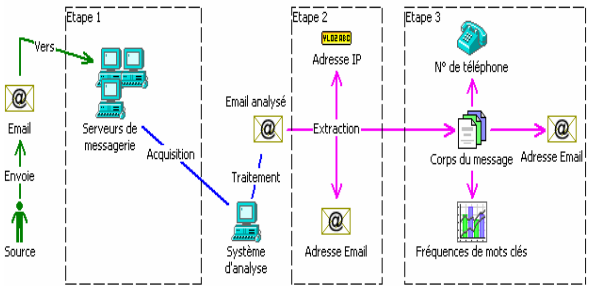

Figure 1 : Description du système par étapes depuis l'envoi d'un e-mail jusqu'aux résultats d'analyse.

Ce projet se découpe en trois étapes :

- **Acquisition des données**
Elle consiste à récupérer les e-mails dans différentes boîtes aux lettres électroniques (alimentées par des organes de police et par les victimes) qui recueillent de tels messages.

- **Intégration et organisation de l'information**
Chaque e-mail est traité afin d'extraire des informations qui seront utiles pour l'analyse, comme le contenu envoyé par l'escroc, les adresses IP, les adresses de l'expéditeur, etc.

- **Analyse de l'information**
Le travail d'analyse porte notamment sur la détection de la langue ou de mots particuliers. Les résultats de ces traitements sont répertoriés dans une base de données pour permettre d'autres analyses globales subséquentes.

La constitution d'un jeu de données permettra de contribuer à la compréhension du phénomène et à la détection d'éventuels réseaux organisés mis en évidence par les liens entre les données extraites (figure 2).

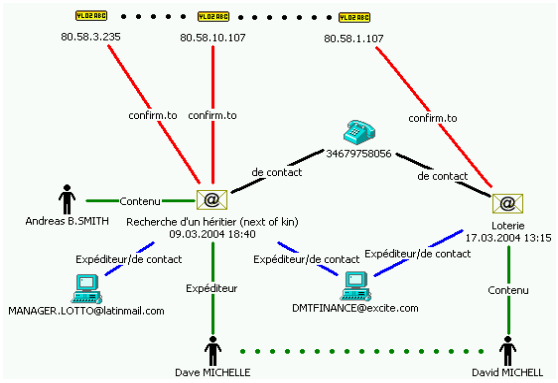

Figure 2 : Les liens obtenus par l'analyse de la forme et du contenu des courriels: des e-mails de type "Nigerian" sont en relation avec des e-mails d'un autre type (promesse de gains à la loterie).

Auteur: **Stéphane Moser**
Répondant externe: **Unil - Institut de Police Scientifique**
Prof. responsable: **M. Christian Buchs**
Sujet proposé par: **Unil -Institut de Police Scientifique**

Hes·so
Haute Ecole Spécialisée
de Suisse occidentale