Chapter 12
SCAMSTOP:
A Platform for Mitigating Fraud in VoIP Environments

Yacine Rebahi
Fraunhofer Fokus, Germany

Reinhard Ruppelt
Fraunhofer Fokus, Germany

Mohamed Nassar
INRIA Research Center Nancy – Grand Est, France

Olivier Festor
INRIA Research Center Nancy – Grand Est, France

ABSTRACT

In traditional telecommunication networks, fraud accounts for significant annual losses at an average up of 5% of the operators' revenue and still increasing. The current shift towards Voice-over-IP (VoIP) networks increases to exposure to fraud due to the lack of strong built-in security mechanisms and the full usage of the open Internet. In this book chapter, the authors discuss an anti-fraud framework they are currently developing within the SCAMSTOP project. Although a short description of the framework is provided, the focus of this chapter is mainly on the methods used to detect fraudulent activity. In particular the authors focus on unsupervised methods including signature and clustering based techniques. Preliminary testing results are also discussed.

INTRODUCTION

The openness, innovative services and low cost structure of Voice-over-IP services has helped providers to attract large numbers of subscribers over the past few years. These same reasons have, unfortunately, also attracted attackers and malicious users who find in these new packet-based networks an opportunity to earn money in a fraudulent way. In traditional telecommunication networks, various experts estimate that fraud accounts for annual losses at an average of 3% to 5% of the operators' revenue. This portion is even

DOI: 10.4018/978-1-4666-1888-6.ch012

still increasing at a rate of more than 10% yearly. Hence, with the openness of the VoIP technology an even higher threat of fraud and higher losses of revenue are expected.

In this chapter, we are interested in investigating the fraud and service misuse problem in VoIP environments. For instance, we study,

- Under which forms does this problem appear in such networks,
- How difficult is VoIP fraud to be detected,
- What can we reuse / adapt from the existing techniques and algorithms in the fight against fraud, and
- Which kind of data do we need to explore to look for fraud patterns?

The mentioned issues and some others are being discussed in the SCAMSTOP project, an FP7 funded collaborative project[1]. The latter aims at designing and implementing innovative and adaptive algorithms for misuse and fraud detection in the VoIP domain. By designing these algorithms, we are not only aiming at achieving a high detection rate but also targeting a scalable architecture ensuring low processing and limited memory usage so as to ensure the applicability of these algorithms in large scale VoIP deployments.

BACKGROUND

Fraud can be defined as any activity that leads to the obtaining of financial advantage or causing of loss by implicit or explicit deception. It is the mechanism through which the fraudster gains an unlawful advantage or causes unlawful loss. This can be as simple as telling a lie to obtain some compensation benefits. Fraud losses keep impacting every business enterprise. The costs of fraud are passed on to the society in the form of increased customer inconvenience, opportunity costs, increased prices for goods and services, and even additional criminal activities funded by the fraudulent gains.

Although Fraud (or scam) is a problem that spans most areas of our daily life (e. g., telecommunications, banking and finance, insurance, e-commerce), we will be more concerned in this document with the fraudulent activities that might occur in VoIP environments.

FRAUD CLASSIFICATION

The classification of fraud can be achieved in different ways according to the point of view from which these activities are observed. However, the categorization that is generally cited in the literature (Bolton & Hand, 2002) is the following,

Subscription fraud: this occurs from obtaining an account or service, often with false identity details, without the intention of paying. The account is usually used for call selling or intensive self-usage.

Superimposed fraud: this type of fraud occurs when a fraudster uses a service or an account without having the necessary authority. In other words, a fraud is said to be superimposed when a fraudster illegally gets resources from legitimate users by gaining access to their phone accounts. The fraudster is said to be an insider when the corresponding activity is committed by an employee of the telecommunication operator. This type of fraud is said to be external if it is committed by a member of the general public. With other respects, this kind of fraud can be detected by the appearance of unknown calls on a bill.

A SHORT STATE OF THE ART

The common techniques used for fraud detection are: rule-based techniques, data-mining as well as both supervised and unsupervised machine learning.