Chapter 26

DETECTING FRAUD IN THE REAL WORLD

Michael H. Cahill
Lucent Technologies, New Providence, NJ 07974, USA
mikecahill@lucent.com

Diane Lambert
Bell Labs, Lucent Technologies, Murray Hill, NJ 07974, USA
dl@bell-labs.com

José C. Pinheiro
Bell Labs, Lucent Technologies, Murray Hill, NJ 07974, USA
jcp@bell-labs.com

Don X. Sun
Bell Labs, Lucent Technologies, Murray Hill, NJ 07974, USA
dxsun@bell-labs.com

Abstract Finding telecommunications fraud in masses of call records is more difficult than finding a needle in a haystack. In the haystack problem, there is only one needle that does not look like hay, the pieces of hay all look similar, and neither the needle nor the hay changes much over time. Fraudulent calls may be rare like needles in haystacks, but they are much more challenging to find. Callers are dissimilar, so calls that look like fraud for one account look like expected behavior for another, while all needles look the same. Moreover, fraud has to be found repeatedly, as fast as fraud calls are placed, the nature of fraud changes over time, the extent of fraud is unknown in advance, and fraud may be spread over more than one type of service. For example, calls placed on a stolen wireless telephone may be charged to a stolen credit card. Finding fraud is like finding a needle in a haystack only in the sense of sifting through masses of data to find something rare. This chapter describes some issues involved in creating tools for building fraud systems.
that are accurate, able to adapt to changing legitimate and fraudulent behavior, and easy to use.

**Keywords:** Customer profiles, Customer relationship management, Dynamic databases, Incremental maintenance, Massive data, Sequential updating, Transaction data, Thresholding.

1. **Introduction**

Fraud is a big business. Calls, credit card numbers, and stolen accounts can be sold on the street for substantial profit. Fraudsters may subscribe to services without intending to pay, perhaps with the intention of re-selling the services, or even the account itself, at a low cost until shut down. Call sell operations may extend their lives by subverting regulatory restrictions that are in place to protect debtors. Gaining access to a telephone or telephone line by physical intrusion still accounts for some fraud. Fraudsters also focus on the people who use and operate the network by applying "social engineering" to instruct an unsuspecting subscriber or operator to unknowingly agree to carry fraudulent traffic. Large profits have justified the growth of a well-organized and well-informed community of fraudsters who are clever and mobile. Fraud is also important to shady organizations that want to communicate without leaving records of their calls that can be traced back to them. Domestically, *Telecom and Network Security Review* (Vol. 4(5), April 1997) estimates that fraud losses in the U.S. telecommunications industry amount to between 4% and 6% of revenue. Internationally, the figures are generally worse, with several new service providers reporting losses over 20%.

Many service providers respond by building fraud control centers. They acquire multimillion dollar network and operations systems, hire and train staff for 24-by-7 operations, educate customers, require the use of Personal Identification Numbers, partner with competitors and law enforcement agencies, perform internal audits, and constantly tune their operations. Automated fraud detection systems may detect calls to certain "hot numbers", simultaneous use of calling cards in distant locations, which is unlikely except in the case of fraud, or other patterns of usage that are known to be associated with fraud. Such efforts have helped to reduce fraud, but the set of fraudsters is continually replenished and fraudsters have been able to continue to operate.

Detecting fraud is hard, so it is not surprising that many fraud systems have serious limitations. Different systems may be needed for different kinds of fraud (calling card fraud, wireless fraud, wireline fraud,