Journal of Parallel and Distributed Computing **61**, 915–925 (2001)
doi:10.1006/jpdc.2000.1720, available online at http://www.idealibrary.com on IDEAL®

An Unsupervised Neural Network Approach to Profiling the Behavior of Mobile Phone Users for Use in Fraud Detection

Peter Burge and John Shawe-Taylor

Department of Computer Science, Royal Holloway University of London, Egham, Surrey TW20 0EX, England
E-mail: BurgeP@logica.com, jst@rhul.ac.uk

Received February 18, 2000; published online May 10, 2001

This paper discusses the current status of research on fraud detection undertaken as part of the European Commission-funded ACTS ASPeCT (Advanced Security for Personal Communications Technologies) project, by Royal Holloway University of London. Using a recurrent neural network technique, we uniformly distribute prototypes over toll tickets, sampled from the U.K. network operator, Vodafone. The prototypes, which continue to adapt to cater for seasonal or long term trends, are used to classify incoming toll tickets to form statistical behavior profiles covering both the short- and the long-term past. We introduce a new decaying technique, which maintains these profiles such that short-term information is updated on a per toll ticket basis whilst the update of the long-term behavior can be delayed and controlled by the user. The new technique ensures that the short-term history updates the long-term history applying an even weighting to each toll ticket. The behavior profiles, maintained as probability distributions, form the input to a differential analysis utilizing a measure known as the Hellinger distance between them as an alarm criterion. Fine tuning the system to minimize the number of false alarms poses a significant task due to the low fraudulent/non-fraudulent activity ratio. We benefit from using unsupervised learning in that no fraudulent examples are required for training. This is very relevant considering the currently secure nature of GSM where fraud scenarios, other than subscription fraud, have yet to manifest themselves. It is the aim of ASPeCT to be prepared for the would-be fraudster for both GSM and UMTS.

1. INTRODUCTION

When a mobile originated phone call is made or various intercall criteria are met the cells or switches that a mobile phone is communicating with produce

0743-7315/01 $35.00

information pertaining to the call attempt. These data records, for billing purposes, are referred to as toll tickets. Toll tickets contain a wealth of information about the call so that charges can be made to the subscriber. By considering well-studied fraud indicators these records can also be used to detect fraudulent activity. By this we mean interrogating a series of recent toll tickets and comparing a function of the various fields with fixed criteria, known as triggers. A trigger, if activated, raises an alert status which cumulatively would lead to investigation by the network operator. Some example fraud indicators are that of a new subscriber making long back-to-back international calls being indicative of direct call selling or short back-to-back calls to a single land number indicating an attack on a PABX system. Sometimes geographical information deduced from the cell sites visited in a call can indicate cloning. This can be detected through setting a velocity trap.

Fixed trigger criteria can be set to catch such extremes of activity, but these absolute usage criteria cannot trap all types of fraud. An alternative approach to the problem is to perform a differential analysis. Here we develop behavior profiles relating to the mobile phone's activity and compare its most recent activities with a longer history of its usage. Techniques can then be derived to determine when the mobile phone's behavior changes significantly. One of the most common indicators of fraud is a significant change in behavior.

The performance expectations of such a system must be of prime concern when developing any fraud detection strategy. To implement a real time fraud detection tool on the Vodafone network in the United Kingdom, it was estimated that, on average, the system would need to be able to process around 38 toll tickets per second. This figure varied with peak and off-peak usage and also had seasonal trends. The distribution of the times that calls are made and the duration of each call is highly skewed.

In this paper we present one of the methods developed under ASPeCT that tackles the problem of skewed distributions and seasonal trends head on using a recurrent neural network technique that is based around unsupervised learning. We envisage this technique forming part of a larger fraud detection suite that also comprises a rule-based fraud detection tool and a neural network fraud detection tool that uses supervised learning on a multilayer perceptron. Each of the systems has its strengths and weaknesses but we anticipate that the hybrid system will combine their strengths.

The following section discusses in more detail the concept of behavior profiling for the purposes of performing a differential analysis. Following this, in Section 3, we present the neural network prototyping technique and the way these prototypes can be used to generate statistical behavior profiles. We then introduce a new technique for maintaining behavior profiles. The short term history is calculated and updated on a per toll ticket basis and thus there is a cumulative decay on the influence of a toll ticket as time passes. In contrast to previous work we update the long-term behavior profile only after a fixed number B of toll tickets. This ensures that altered behavior has a marked effect on the short-term history before having any influence on the long-term behavior profile and hence maximizes the chances of detecting anomalies. A new decaying method enables us to update the long-term