A Model for the Visual Data Mining of Call Patterns

Work in progress paper

Stephanus Francois du Toit, Andre Calitz
Department of Computer Science and Information Systems
Nelson Mandela Metropolitan University, PO Box 77000, Port Elizabeth, 6031
Tel: (041) 504 2080, Fax: (041) 504 2831
Email: Stephanus.duToit@nmmu.ac.za, Andre.Calitz@nmmu.ac.za

Abstract – The number of call centres in South Africa have increased the past number of years. Large amounts of data are generated by calls received at the calls centres. Data mining techniques are not currently utilised on the call event data. This paper proposes the application of data mining for useful information extraction at call centres on call event data. Furthermore the use of effective visualisation techniques of the data is proposed in order to make informed managerial decisions.

Index Terms – call centre data mining, visual data mining model

I. Introduction

The South African Call centre industry is fast becoming one of the world leaders. An independent analyst, Datamonitor, has predicted an increase from 494 call centres in South Africa in 2003 to 939 in 2008.[8]

Call centres are classified according to their purpose. Inbound call centres, call centres with only inbound calls/traffic, customer generated calls, generally used for after sales services. Outbound call centres, traffic is representative generated, e.g. telemarketers, and call centres that have both incoming and outgoing calls[12].

Merchants, a subsidiary of Dimension Data, is a South African company that provides call centres to companies seeking to provide services through a call centre. They provide complete solutions, the location, hardware and the agents. Clients include South African cellular company Vodacom, and America online. Merchants receive approximately 1.2 million phone calls a month to the various call centres provided to customers.

II. Data Mining of Call Patterns in Telecommunications

In telecommunications research is conducted on the data mining of call patterns[10, 11]. Some of the current applications are fraud detection[7, 6, 1], classification of residential, industrial, business areas and telemarketing[10].

Data mining for call patterns in telecommunications is achieved through performing data mining operations on call event data. Call event data is very course and does not contain useful information at a customer level. An example of call event data is [10]:

- Originating number;
- Terminating number;
- Date and Time;
- And Call Duration.

The originating number is the number that initiates a call, terminating number is the number that receives a call, date and time is the date, and time that a call starts. Duration is the duration of the call.

In order to gain knowledge from this data at a customer level, it needs to be processed before knowledge acquisition is attempted. An example of the processed data could be as follows for a specific time frame P[10]

- Average call duration for P;
- % unanswered calls for P;
- % calls originating or terminating in different area code for P;
- % weekday calls for P;
- % daytime calls for P;
- Average number of received calls per day for P;
- And average number of dialled calls per day for P.

This data is derived form the call event data and is ready for data mining to produce results at a customer level as in apposed to a call level.

III. Data Mining in Call Centres

Data mining is employed in call centres with the goal of improving Customer Relationship Management (CRM)[9].

Outbound call centres utilise predictive diallers that attempt to dial a potential customer whilst a representative is not yet available, attempting to connect them as soon as the representative is available reducing a telemarketers idle time[12], [10]. Merchants, currently do not perform data mining on its call event data. Statistical reports of call volume versus time is the only knowledge gained from call data and is used to predict the amounts of agents that will be required during the day at various times.

IV. Data Mining of Call Patterns in Call Centres

The data captured with regards to the call event in the Merchants call event databases is stored centrally. A single table containing all calls made to the various clients call centres. Currently the data for the past two years, since they started storing call event data is being stored. There are potential benefits in data mining at a customer level and at call event level.

V. Visualisation of Data Mining

Gaining knowledge from data is a timely and costly exercise and this has been described as the knowledge acquisition bottleneck[3]. Visual data mining has proven useful in aiding with discovering hidden knowledge in large datasets through data mining.

Two types of visualisation has been described in H. L. Maria Cristina Ferreira de Oliveira [5] for visual data mining. Visualisation of the data, this is visualisation of large data sets, and visualising the data mining results itself. The second visualisation technique is visualisation of data mining.

A three orthogonal approach to visualisation of large data sets has been described in previous literature[4, 1, 2]. The first aspect is visualisation this can be described as the physical representation of the data to users. The second is interaction, this is allows the user to explore the visual representation in order to gain more insight into the knowledge in the data. The third is distortion, which is used to allow the user to focus on certain parts of information, without losing perspective on the overall data.

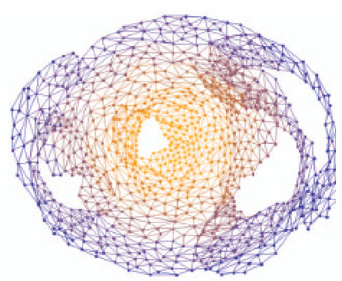

Figure 1 – Topological fisheye view

Figure 1 is a topological fisheye view. It illustrates how certain data can be highlighted without loosing overall perspective on the data.
Visualisation of data mining refers to the role of visual feedback of the data mining algorithm itself. Visualisation can aid in the better understanding of the results of an algorithm by the stakeholders involved and also provide insight into how the algorithm works. Understanding data mining algorithms are of benefit to the end users involved with the data mining results, and give them more control in the knowledge discovery process.

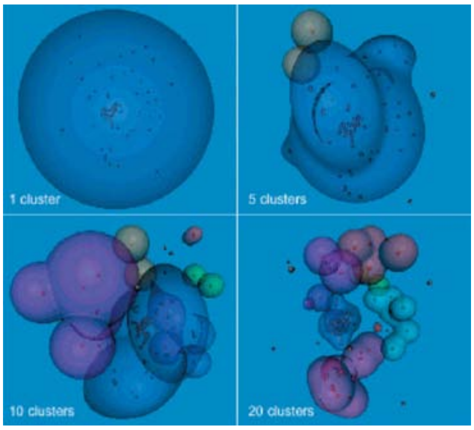

Figure 2 – visualisation of clustering of data in a large data set

Figure 2 (from [5]) shows a visualisation of clusters in a data set.

VI. Conclusion and future work

Visual Data mining has been deployed and aids in knowledge discovery in databases. Call centres have large amounts of data that is currently only used for statistical purposes. The author proposes that a model and prototype for visual data mining of call patterns to test the value of such a system for call centres. Future investigations could include mining not only on call event data but on a combination of call event data and representative captured data that describes the type of problem a phone call was about. This model would allow managers to make more informed decisions.

References

[1] E. E. K. Daniel A. Keim, Stephen C. North, *Visual Exploration of Large Telecommunication Data Sets*, User Interfaces to Data Intensive Systems (1999), pp. 12.

[2] H.-P. K. Daniel A. Keim, *Visualization Techniques for Mining Large Databases: A Comparison*, IEEE Transactions on Knowledge and Data Engineering, 8 (1996), pp. 923-938.

[3] F. David, C. H. John, K. Russell, A. M. Heidar, S. Charles, S. Michael and T. Efraim, *The impact of machine learning on expert systems*, *Proceedings of the 1993 ACM conference on Computer science*, ACM Press, Indianapolis, Indiana, United States, 1993.

[4] N. J. R. T. F. P. Eleftherios E. Koutsofios Stephen C. North Florham Park, N.J Daniel A. Keim, *Visualizing large-scale telecommunication networks and services (case study)*, *Proceedings of the Conference on Visualization '99: Celebrating Ten Years*, IEEE Computer Society Press Los Alamitos, CA, USA, San Francisco, California, United States, 1999, pp. 457-461.

[5] H. L. Maria Cristina Ferreira de Oliveira, *From Visual Data Exploration to Visual Data Mining: A Survey*, IEEE TRANSACTIONS ON VISUALIZATION AND COMPUTER GRAPHICS, 9 (2003), pp. 378-394.

[6] D. L. Michael H. Cahill, Jos´e C. Pinheiro, Don X. Sun, *Detecting fraud in the real world*, in P. M. P. James Abello, Mauricio G. C. Resende, ed., *The Handbook of massive Datasets*, Kluwer Academic Publishers, 2001, pp. 911 - 929.

[7] C. M. H. a. C. A. P. Pablo A. Estévez, *Subscription Fraud Prevention in Telecommunications using Fuzzy Rules and Neural Networks*, Expert Systems with Applications, 31 (2006).

[8] S. i. reporter, *SA call centres score again*, SouthAfrica.info, 2006.

[9] I. J. Stephen White, *Investigation into the Application of Data Mining Techniques to Classification of Call Centre Data*, *The 2004 IFIP International Conference on Decision Support Systems*, PRATO, TUSCANY, ITALY, 2004.

[10] G. M. Weiss, *Data Mining in Telecommunications*, in O. M. a. L. Rokach, ed., *Data Mining and Knowledge Discovery Handbook: A Complete Guide for Practitioners and Researchers*, Kluwer Academic, 2005, pp. 1189-1201.

[11] G. M. Weiss, *Mining with rare cases*, in O. M. a. L. Rokach, ed., *Data Mining and Knowledge Discovery Handbook: A Complete Guide for Practitioners and Researchers*, Kluwer Academic, 2005, pp. 765-776.

[12] W. Whitt, *Stochastic models for the design and management of cutomer contact centers: some research directions*, Columbia University, 2002.

Stephanus Francois du Toit received is B.Sc. in computer science and applied math from the University of Port Elizabeth (currently Nelson Mandela Metropolitan University) in 2002 and his B.Sc. hons in 2003. Currently he is studying his M.Sc. at the Nelson Mandela Metropolitan University.