Knowledge Discovery in Integrated Call Centers: A Framework for Effective Customer-Driven Marketing

Paul Xia

EIS International Inc.
1351 Washington Blvd., Stamford, CT 06902
pxia%EIS_International@mcimail.com

ABSTRACT

As call centers become more pervasive, the customers seek individualized service and greater attention. The call centers are becoming the contact centers - a one-stop, single interface for all interactions - from pre-sales to post-sales, and continuing relationship. This paper presents some results in utilizing data mining in managing customer profile toward a greater business advantage.

One solution is to build a predictive customer profile based on the customer's Lifetime Value (LTV). Many different types of data mining techniques can be combined to meet this challenge.

This paper explores the solution to using knowledge discovery methods in integrated inbound/outbound call center environments. Significant performance gain is reported on a customized knowledge acquisition system over a conventional approach.

KEYWORDS

Call Center; Data Mining; Knowledge Discovery; Lifetime Value

1. INTRODUCTION

In a typical integrated call center environment, operators either answer calls from callers or dial customers for diverse purposes. In either situation, the centers want the best utilization of resources including human operators (hereafter called agents), computing resources, trunk lines, other media (such as Voice Response and Web Page), and the physical facilities. Operational productivity is maximized by minimizing the number of calls involving agents without sacrificing customer satisfaction.

A customer places a call or accepts offers/services for different reasons. By the same token, different factors result in their satisfaction and desire to call again or to buy more. The challenge, therefore, is to identify shared characteristics of customers (combinations of behaviors, attitudes, lifestyles, and demographics) and find out what differentiates them from other groups of customers. Once customers are truly understood on an individual basis, a more effective outbound campaign can be launched. Similarly, resources can be allocated to calls from your best customers and pay the greatest attention to their individual needs.

2. CUSTOMER-DRIVEN APPROACH

The goal of a customer-driven approach is to determine the average customer's Lifetime Value (LTV) in a fairly accurate way so that customer LTV models can be constructed. By classifying customers into different groups with different LTVs, a more efficient call center communication can be organized.

A customer-driven approach usually includes the following major components: (See Fig. 1)

1. Identify customers
2. Differentiate each customer
3. Interact with each customer
4. Customize products for each customer

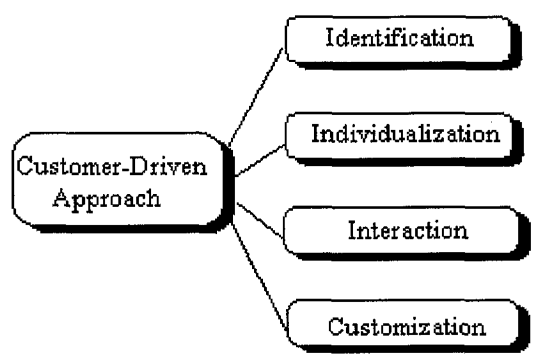

Figure 1: Main Components of Customer-Driven Marketing

3. DATA MINING: A TECHNIQUE FOR CONSTRUCTING AN LTV MODEL

In creating the dynamic LTV model defined in Section 2, the following steps can be taken:

1. Build a customer profile
2. Create an LTV base model

3. Create an In/Outbound call campaign with forecasting capacity
4. Create an In/Outbound call survey with learning capacity
5. Identify warehouse data and external supplements
6. Repeat step (3), (4), (5) to generate a more complex and accurate LTV model

The following sub sections discuss how data mining technology can be applied in each step.

3.1 Build a Customer Profile

Customer profile building is a process of segmentation according to definite types of customer behavior.

If similarity measures and segmentation examples are known, clustering can be achieved with techniques such as k-nearest neighbors. If no classification is known, but some similarity exists between customers, clusters are formed maximizing the degree of similarity between callers in the same partition. Depending upon how and when the new classification request has been made, different approaches can be taken.

For building the primary customer segmentation, or if new customers have been acquired from external sources, the global similarity could be maximized with Genetic Algorithm (GA).

The similarity in our interactive KDD system is pre-defined by the representation based on a customer's profile and a call center's activities such as telemarketing, donation, collection, reservation, and help-desk.

3.2 Create an LTV Base Model

The most important step in the LTV modeling process is to build time series functions that measure some properties of partitions as functions of certain parameters of the partition.

Suppose a time period A is given, and suppose A can be divided into disjointed set of time periods called Aj. One can define a function V(Aj) that measures some property for that time period. Then one can build a global function V(A) as the sum of V(Aj) to represent LTV:

$$LTV = V(A) = \sum_{j} V(Aj) \qquad (1)$$

Since a customer's LTV is meaningful only when it is associated with a particular service or product, and one can define a function S(Mk, Aj) to correspond to the value of the service/product Mk in time period Aj for a particular customer, V(Aj) can be expressed as:

$$V(Aj) = \sum_{k} S(Mk, Aj) \qquad (2)$$

The computation in (2) is facilitated by a judicious choice of a vector comprising of elements (Mk, Aj).

Customer value can be quantified in various levels and in various time intervals. The simplest characterization is high and low in value, and time as present and future.

3.3 Create an Inbound/Outbound Call Campaign with Forecasting Capacity

Using a caller's personal profile, effective caller segmentation can be created.

Induction trees are being used to extract prediction rules for the future based on past data, in order to find groups of people with high similarity in an LTV model.

When an inbound/outbound call succeeds, if the product or service is accepted, basket analysis would be conducted, and other associated merchandise items generated from it are also offered.

3.4 Create an Inbound/Outbound Call Survey with Learning Capacity

Because of the contact interface (web or/and phone) and its nature (such as quick response, single/multiple selection(s), sophistication of users, etc.), the selected data mining technology must be:

- fast
- understandable
- easily translated to a relational database
- user friendly

A modified C4.5-like algorithm was developed to meet above criteria for supervised classification.

If an instance can not be found in real time, unsupervised clustering is involved and more constraints are imposed.

3.5 Identify External/Warehouse Data Supplements

Our interactive KDD system provides data warehouse and host interface and a set of utilities for the noise handling.

4. INBOUND/OUTBOUND CALL PROCESSING MODEL

The following figure (Fig. 2) demonstrates simplified call processes with knowledge learning capacities.

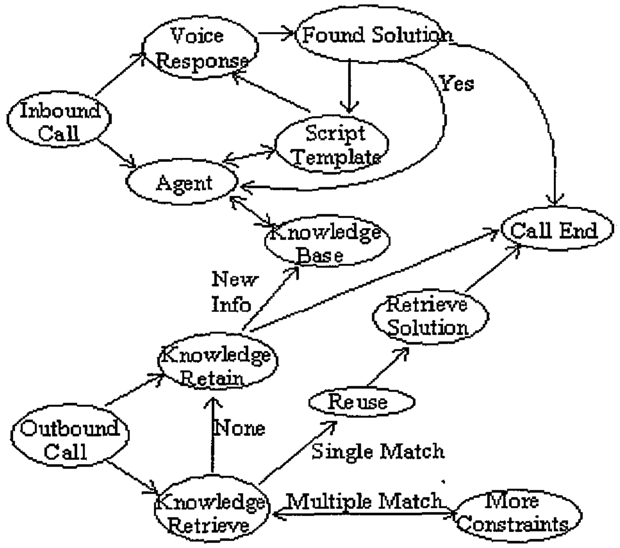

Figure 2: Call Process with Knowledge Learning

5. MODEL SYSTEM IMPLEMENTATION

5.1 Description

The following flowchart (Fig. 3) demonstrates the basic information flow of a call center Knowledge Acquisition system.

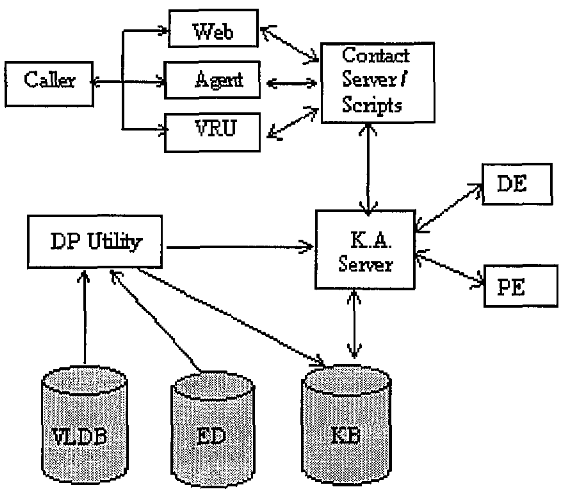

Figure 3: Simplified Interactive Learning Database System

VRU: Voice Response Unit
Scripts: Information Container
K.A.Server: Knowledge Acquisition Server
DE: Discovery Engine
PE: Predictive Engine
DP Utility: Data Processing Utility including data segmentation, transformation and aggregation
VLDB: Very Large DB, Warehouse Data
E.D: External Data
KB: Knowledge Base/Mine Data

By talking with the callers via the telephone, IVR, and/or the Internet, the system learns a lot about their individual needs, satisfaction, demographic and psychological characteristics. All the information from callers is analyzed on fly. The key information regarding what a customer says and wants can be fed into a central database, and a more accurate LTV model can be constructed.

The database learning system presented above provides a mechanism for:

- Translation of processed database information (from data warehouse and external data sources) into a form suitable for use by discovery and prediction engines
- Extraction of rules and other knowledge from the database, which is stored as meta objects in a knowledge base
- Dynamic agent scripts created on the fly to understand and solve caller's needs

5.2 Implementation

The knowledge base including mined data is stored and constructed as meta objects under ODBMS for:

- performance gain
- intelligent and efficient storage
- flexibility in constraint processing for various types of data
- better representation of the regular knowledge of attributes

5.3 Performance

The speed of finding similarity and rules is critical for caller access. Many approaches are used to meet this timing requirement. For instance, the modified genetic algorithms used as a filter to reduce the number of features needed to describe an instance.

5.4 Data/Knowledge Representation

The knowledge representation is created by a dynamic memory model which organizes similar cases under general categories, which includes:

- A typical case for that time period for the model LTV
- Common features
- Contributing or discriminating features
- Relevance weight associated with contributing or discriminating features

6. COMPARISON OF KDD MODEL TO CONVENTIONAL APPROACHES

Two data sets were used for model development and testing: a small demographics data set of approximately 100 MB containing 600 attributes on 35,000 callers and buying-pattern information, and a larger demographics data set of approximately 1.2 GB containing 70 attributes on 3,000,000 callers. Data were obtained from many sources. Based on simulation and some field tests, we found significant performance gain as depicted in Figures 4 and 5:

- The sales close rate improved dramatically by at least 30%. The amount of improvement depended upon how well the caller's needs were addressed, proper products, etc.

- Dramatic reduction of returned merchandise. It was achieved by tailoring products/services more tightly into each customer's unique needs as well as early fraud detection.

- Because of the overall improvement of customer satisfaction, caller/prospects come back far more often than before.

- The customer's satisfaction has been enhanced by 15% or more due to the understanding of their diversified needs and the use of individualized scripts which provides much better service.

- Since customer profiles and trends were measured dynamically, the agent can cross-sell products/ services to customers. Therefore, not only did the revenue per sale increase by 150%, but also the cost per lead was down by 70% or more.

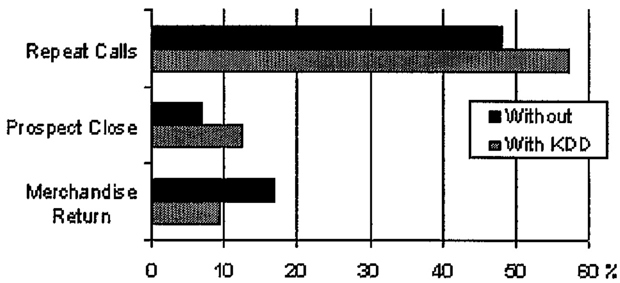

Figure 4: Return/Close/Repeat Business Comparison

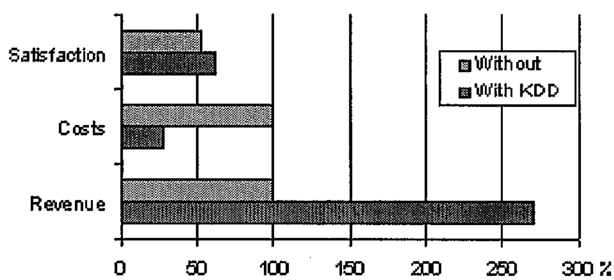

Figure 5: Revenue/Costs/Satisfaction Comparison

7. CONCLUSION

The application of interactive-knowledge discovery to a call center environment is a valuable approach. It maximizes agent productivity while meeting desired performance requirements. It beats the traditional mining process by over 30% or more in every measure. Its flexibility in adjusting to any customer's condition change makes it ideally suited to dynamic conditions in both inbound and outbound calling.

Much can be gained by positioning a call center as an interactive medium. A business entity can acquire critical customer knowledge, recognize consumers' diverse needs, respond to these needs individually and serve them. Service itself can now be variable, determined by a current and future profitability, which is combined with company's goals for increasing customer acquisition, referrals, repeat purchases and cross-sales.

8. BIBLIOGRAPHY

[1] Michie, D. ,Spiegelhalter and Taylor, **Machine Learning, Neural and Statistical Classification**, Ellis Horwood, 1994

[2] Fayyad, U., Piatetsky-Shapiro, Smyth, and Uthurusamy, **Advances in Knowledge Discovery and Data Mining,** MIT Press, 1996

[3] Lawerence Davis, **Handbook of Genetic Algorithms**, Van Nostrand Reinhold, 1991

[4] Quinlan, J. R., **C4.5: Programs for Machine Learning,** Morgan Kaufmann, 1992

[5] Don Peppers, Martha Rogers, **Enterprise One To One**, Doubleday, 1997

[6] **Proceedings of the 1st International Conference On Knowledge Discovery & Data Mining,** AAAI Press, 1995

[7] **Proceedings of the 2nd International Conference On Knowledge Discovery & Data Mining,** AAAI Press, 1996