# A Framework for Call Center Decongestion Using Sequential Pattern Analysis

Eugene Rex L. Jalao
Department of Industrial Engineering and Operations Research
University of the Philippines Diliman
Quezon City, Philippines
e-mail: eljalao@up.edu.ph

***Abstract*— In an effort to improve its customer service, mobile telecommunication companies implemented various customer service channels like call center hotlines, text messaging, email or web self-service where subscribers conduct various after-sales transactions. Nowadays, a hotline call center via a customer service representative is the top choice alternative preferred by subscribers. However, it is noted that the cost of each call when transacting on a hotline is much greater than the cost of the other channels. Furthermore, subscribers get easily irritated when they need to wait for a long time to avail the services. In order to address the problem of reducing hotline calls, as well as to reduce the cost of customer service transactions, subscriber call transactions are analyzed in this paper to predict the next type of call that the subscriber will transact. A sequential pattern analysis methodology is applied and frequent sequences of calls are collected. Given the frequent sequences, the sequences of transaction calls are identified and a corresponding campaign is introduced to intercept the new calls and divert the transaction to less costly customer service channels.**

**Keywords:** ***call center; decongestion; hotline; sequential pattern analysis;***

## I. INTRODUCTION

Delivering good customer service particularly in the telecommunications or telco industry is a must to keep its customers from churning. In this industry, customer service is captured from pre-sales research, to in-life transactions and finally to the renewal of the plan through the procurement of new or upgraded handsets. Basically the entire subscriber life cycle is covered. Hence, to stay with the subscribers every step of the way, these telcos implemented various traditional customer service channels, e.g., physical retail stores, fax numbers, email and contact center hotlines. Subscribers would just transact their needs in one of these channels through the aid of customer service representatives (CSRs).

However, these traditional channels are expensive to maintain and operate. Operating costs include: labor costs of CSRs, rent, and other overhead expenses. In the Philippines, given that the average call and wrap up time metric from the International Finance Corporation is 6 minutes [1], and operating costs of a call center hover from $8 – $14 per hour [2], an average call would cost the telco company $0.8 – $1.4 per customer transaction. If the hotline receives 100,000 calls per month, this would lead to approximately $80,000 to $140,000 operating costs for the telco.

Hence, in an effort to reduce transaction costs, telcos implemented alternative self-service channels to divert subscribers from using the aforementioned traditional costly channels. Some examples of these alternative channels would be: web applications, SMS, and smartphone apps self-service channels. Some companies even utilize social media through official pages to reach their customers. Yet, not all transactions can be deployed in all channels. For example, disputes on billing can only be addressed when talking to a CSR. On the other hand, simple transactions like bill balance inquiry can be done on any channel. The paper by Kumar and Telang [3], presented that the cost of transacting through self-service channels (< 1$ per transaction) is way cheaper than traditional channels with CSRs (between $5 and $10).

On the other hand, a survey by McCarthy and Giles [4] asserts that a phone call with a CSR is the top preferred channel by approximately 75% of the respondents. Additionally, the top most frustrating thing about customer service is the long waiting time to speak to a CSR. Therefore, it is prudent for telcos to keep operating these call centers but on the other hand, having a congested hotline would also lead to negative impacts to their customer service. In essence, telcos would want transactions that can be transacted through self-service channels be diverted from these costly customer service channels.

Given this background, this paper attempts to profile telecommunication subscribers in an effort to decongest call center hotlines and divert possible self-service transactions to other low cost self-service channels. This paper is further divided as follows: Section II provides an overview of current published methodologies, Section III presents the framework of the model while Section VI shows the results and discussion when applied to a local telecommunications company. Section V provides the conclusions and further research.

## II. REVIEW OF RELATED LITERATURE

There have been several attempts to decongest a call center hotline through various predictive and prescriptive analytics methodologies.

Forecasting methodologies are considered in [5] - [8]. These models specifically forecast the volume of calls that arrive at a single time interval. These are mainly for

 ISBN: 978-1-61208-510-4

operational issues and mainly address the issue of how much work they have to do.

A lot of call centers deal with highly erratic demand that is also time-based. The time-based component is relatively easy to handle by adjusting agent staffing [9]. Some examples would be in [10], where a proposed framework that combines linear programming with simulation to recommend a schedule. Search methods are developed which used queueing theory to produce agent schedules for a multi-skill call center is done in [11]. However, it's the random component that contributes to the complexity of the demand.

Other models tried to reduce the call demand by limiting calls admitted to the hotline. A paper by Omeci et al. [12] proposed a selective form of call admission by selectively admitting calls according to their relative importance to the organization.

Given these solutions, a methodology that could predict the most likely next transaction call type of the subscribers would be beneficial such that low priority transactions can be routed to other low cost channels. A methodology that fits this bill is Sequential Pattern Analysis (SPA). Agrawal and Srikant [13] defined SPA as a methodology to extract frequent sequences within a set of transactions. There have been several applications that used the SPA algorithm: Choi et al. in [14] applied SPA on online-product recommendation systems, Huang et al. in [15] proposed a knowledge-assisted sequential pattern analysis framework to identify the patterns of the uterine contractions as well as labor contractions. Chen et al. [16] in proposed to use SPA to forecast potential customers by identifying attributes with high level of association.

Given the multitude of applications of SPA, this paper hypothesizes that the use of SPA in determining the frequent sequences of calls of subscribers on a hotline can be done to address hotline decongestion.

## III. PROPOSED FRAMEWORK

The proposed hotline decongestion framework is composed of three components. These are defined as: (1) the Preprocessing Component, (2) SPA Component and the (3) Campaigns Design Component. This section provides an overview of the different components as follows.

### *A. Preprocessing Component*

Call transaction data is extracted from the call database with the following dataset structure as presented in Fig. 1:

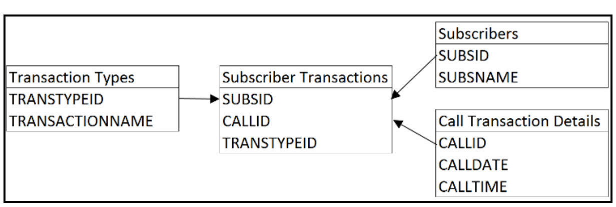


Figure 1. Dimensional Model of required data

The proposed dataset for the SPA component would require to have a granularity of: one row per call with corresponding subscriber and transaction type ID. Hence the following dataset, that joined multiple tables are presented in Fig. 2 as follows:

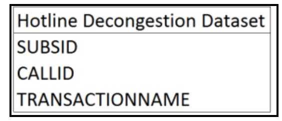


Figure 2. Proposed Dataset Required

Invalid transactions like prank calls, dropped and abandoned calls are likewise eliminated from the dataset leaving only valid transactions. The clean dataset is then sent to the SPA component of the framework.

### *B. SPA Component*

In this component, the SPADE (Sequential PAttern Discovery using Equivalence) algorithm proposed by Zaki [17] is utilized on a R Platform. The package "arulesSequences" is utilized as the main library to compute for the frequent subsequences.

The output of this component is a set of frequent sequence of transactions with format as presented in equation (1).

$$<\{\mathrm{transA}\},\{\mathrm{transB}\},\{\mathrm{transC}\},>\ \textit{support} = s \qquad (1)$$

### *C. Campaigns Design Component*

Given the result of the SPADE algorithm, corresponding campaigns need to be designed to identify the types of future calls of the subscribers. If the type of call can be done in other low-cost channels, a system is in place to intercept that call and divert that transaction.

To illustrate given the frequent transaction in equation (1), if the subscriber called for {transA} then subsequently called for {transB}, and if {transC} can be done in other channels, a text message from the network can be sent to the subscriber to avail of {transC} in other self-care channels instead of calling the hotline. This would hopefully encourage the subscriber to avail of {transC} in other channels, thus reducing calls in the hotline.

## IV. RESULTS AND DISCUSSION

To test the framework, hotline transactions from a leading telecommunications company here in the Philippines are considered and extracted for preprocessing. From the third quarter of 2015, a total of 245,162 calls are extracted from the hotline data warehouse coming from 122,462 unique subscribers. Of the total calls, 23,812 calls are considered invalid since these are invalid transactions and thus are eliminated from the dataset. A snapshot of the final dataset is shown in Fig. 3:

 ISBN: 978-1-61208-510-4 

| SUBID | TRANSID | TRANS |
|---|---|---|
| 10454567890 | 688936812 | FOLLOW UP WITHIN SLA |
| 10454567890 | 688938242 | SALES LEAD |
| 10454567890 | 688939685 | SALES LEAD |
| 10454567890 | 688943163 | REFERRED TO OTHER HOTLINE |
| 10454567890 | 688943751 | FOLLOW UP WITHIN SLA |
| 10454567890 | 688945660 | MECHANICS PROCEDURE |
| 10454567890 | 688945718 | SIM RELATED |

Figure 3. Dataset Preview of the Data

This dataset is fed into the SPADE algorithm with 1% minimum support setting. Results from the dataset showed that there are 22 sequences of length one transaction and 13 sequences of length two. The 13 sequences are presented in Fig. 4.

```
<{"DEVICE CONFIGURATION"},{"SUCCESSFUL NOT INTERESTED"}> 0.02634287
    <{"DEVICE CONFIGURATION"},{"SUCCESSFUL INTERESTED"}> 0.04060852
     <{"MECHANICS PROCEDURE"},{"SUCCESSFUL INTERESTED"}> 0.01628260
           <{"BILLING INQUIRY"},{"MECHANICS PROCEDURE"}> 0.01084418
      <{"DEVICE CONFIGURATION"},{"MECHANICS PROCEDURE"}> 0.01028891
       <{"MECHANICS PROCEDURE"},{"MECHANICS PROCEDURE"}> 0.01476376
     <{"DEVICE CONFIGURATION"},{"DEVICE CONFIGURATION"}> 0.02413810
    <{"SUCCESSFUL INTERESTED"},{"DEVICE CONFIGURATION"}> 0.01192207
               <{"ACCOUNT DETAILS"},{"BILLING INQUIRY"}> 0.01221603
               <{"BILLING INQUIRY"},{"BILLING INQUIRY"}> 0.02843331
           <{"AFTERSALES REQUEST"},{"AFTERSALES REQUEST"}> 0.01172609
              <{"BILLING INQUIRY"},{"AFTERSALES REQUEST"}> 0.01557218
               <{"BILLING INQUIRY"},{"ACCOUNT DETAILS"}> 0.01315510
```

Figure 4. The 13 frequent sequences of length 2.

Given the results from figure 4, we examine the sequential rule in equation (2) as follows:

$$< \{\text{Account Details}\}, \{\text{Billing Inquiry}\}> \textit{support} = 1.2\% \qquad (2)$$

This can be interpreted as: 1.2% of 122,462 or 1470 unique subscribers can be intercepted on the transaction "Billing Inquiry" if they called for "Account Details" beforehand. This could lead to a reduction of at least 1470 calls within a given quarter. The other 12 rules can be designed to maximize the use of the framework to further reduce the number of repeat calls.

## V. CONCLUSION AND FUTURE STUDIES

The hotline decongestion framework, composed: (1) preprocessing, (2) sequential pattern analysis, and (3) campaigns design components, is able to profile subscribers in terms of the call type sequences. Frequent call sequences can be extracted and corresponding campaigns can be developed to intercept and divert call transactions to alternative lower cost self-service channels.

Further improvement of the framework is considered in terms of analyzing the inter-arrival time of the calls. The framework currently studies the sequence of calls but not the amount of time in between calls made by the subscriber.

## REFERENCES


[1] IFC, "Call Center Pricing," 2016. [Online]. Available: http://www.ifc.org/wps/wcm/connect/75ce96004cf85d4f8752c7f81ee631cc/Tool+9.4.+Measuring+Call+Center+Performance.pdf?MOD=AJPERES. pp. 1 [Accessed: 06-Jul-2016].

[2] World Wide Call Centers, "Call Center Pricing," 2016. [Online]. Available: https://www.worldwidecallcenters.com/call-center-pricing/. pp. 1 [Accessed: 06-Jul-2016].

[3] A. Kumar, and R. Telang, "Does the web reduce customer service cost? Empirical evidence from a call center," *Inf. Syst. Res.*, vol. 23, no. 3 PART 1, pp. 721–737, 2012.

[4] C. Mccarthy, M. Giles, W. W. W. O. Com, and C. Mccarthy, "Customer Service : Where are Telcos Investing ? About the authors," *Ovum Signatue Res.*, no. pp 1-2, November, 2011.

[5] J. Taylor, "Short-term electricity demand forecasting using double seasonal exponential smoothing," *J. Oper. Res. Soc.*, vol. 54, no. 8, pp. 799–805, 2003.

[6] R. Soyer and M. M. Tarimcilar, "Modeling and Analysis of Call Center Arrival Data: A Bayesian Approach," *Manage. Sci.*, vol. 54, no. 2, pp. 266–278, 2007.

[7] J. Weinberg, L. D. Brown, J. R. Stroud, J. W. Einberg, L. D. B. Rown, and J. R. S. Troud, "Bayesian Forecasting of an Inhomogeneous Poisson Process With Applications to Call Center Data," *J. Am. Stat. Assoc.*, vol. 102, no. June, pp. 1187–1199, 2007.

[8] H. Shen, and J. Z. Huang, "Interday Forecasting and Intraday Updating of Call Center Arrivals," *Manuf. Serv. Oper. Manag.*, vol. 10, no. 3, pp. 391–410, 2008.

[9] Z. Aksin, M. Armony, and V. Mehrotra, "Customer Behavior Modeling in Revenue Management and Auctions: A Review and New Research Opportunities," *Prod. Oper. Manag.*, vol. 16, no. 6, pp. 665– 688, 2007.

[10] M. T. Cezik, and P. L'Ecuyer, "Staffing Multiskill Call Centers via Linear Programming and Simulation," *Manage. Sci.*, vol. 54, no. 2, pp. 310–323, 2008.

[11] A. N. Avramidis, A. Deslauriers, and P. L'Ecuyer, "Modeling Daily Arrivals to a Telephone Call Center," *Manage. Sci.*, vol. 50, no. 7, pp. 896–908, 2004.

[12] H. E. Ormeci, E. L., and A. N. Burnetas, "Admission policies for a two class loss system with random rewards," *IIE Trans.* \, vol. 34, no. 9, pp. 813–822., 2002.

[13] S. Agrawal, and Rakesh; Ramakrishnan, "Mining sequential patterns," *Proc. Elev. Int. Conf. Data Eng.*, pp. 3–14, 1995.

[14] K. Choi, D. Yoo, G. Kim, and Y. Suh, "A hybrid online-product recommendation system: Combining implicit rating-based collaborative filtering and sequential pattern analysis," *Electron. Commer. Res.*

*Appl.*, vol. 11, no. 4, pp. 309–317, 2012.
[15] Z. Huang, M. L. Shyu, J. M. Tien, M. M. Vigoda, and D. J. Birnbach, “Prediction of uterine contractions using knowledge-assisted sequential pattern analysis,” *IEEE Trans. Biomed. Eng.*, vol. 60, no. 5, pp. 1290–1297, 2013.
[16] W. C. Chen, C. C. Hsu, and J. N. Hsu, “Optimal selection of potential customer range through the union sequential pattern by using a response model,” *Expert Syst. Appl.*, vol. 38, no. 6, pp. 7451–7461, 2011.
[17] M. J. Zaki, “SPADE: An efficient algorithm for mining frequent sequences,” *Mach. Learn.*, vol. 42, no. 1–2, pp. 31–60, 2001.

 **ISBN: 978-1-61208-510-4**