Phone Call Log as a Context Source to Modeling Individual User Behavior

Iqbal H. Sarker
School of Software and Electrical Engineering,
Swinburne University of Technology,
Melbourne, Australia.
msarker@swin.edu.au

Alan Colman
School of Software and Electrical Engineering,
Swinburne University of Technology,
Melbourne, Australia.
acolman@swin.edu.au

Muhammad Ashad Kabir
School of Computing and Mathematics,
Charles Sturt University,
NSW, Australia.
akabir@csu.edu.au

Jun Han
School of Software and Electrical Engineering,
Swinburne University of Technology,
Melbourne, Australia.
jhan@swin.edu.au

Ubicomp/ISWC'16 Adjunct, September 12-16, 2016, Heidelberg, Germany
© 2016 ACM. ISBN 978-1-4503-4462-3/16/09…$15.00
DOI: http://dx.doi.org/10.1145/2968219.2971592

Abstract
Mobile phone can record various types of context data related to a user's phone call activities in its call log. Call log provides temporal context to modeling individual user's phone call response behavior, i.e., when a user accepts, rejects or misses an incoming call. In this paper, we explore the potentiality of phone call log as a context source to modeling call response behavior of individual mobile users. Towards this, we present our initial work to generating temporal rules that capture the user's dominant call response behavior at various times of the day and days of the week, utilizing phone call log. Our preliminary experimental results on real datasets show that context information in call log can be used to model individual's phone call response behavior with high accuracy.

Author Keywords
Mobile Phone, user behavior modeling, call log, temporal context, rule discovery

ACM Classification Keywords
H.3.4 [Systems and Software] Current awareness systems, User profiles and alert services

Introduction
Mobile phones have become increasingly ubiquitous and powerful. Their ability to log call activities information

call-1-raw-data
Time: 12:10:20
Date: 12/04/2015
Call Type: INCOMING
Call Duration: 0
call-2-raw-data
Time: 12:30:20
Date: 13/04/2015
Call Type: INCOMING
Call Duration: 125
call-3-raw-data
Time: 12:50:20
Date: 14/04/2015
Call Type: MISSED
Call Duration: 0

Figure 1: Sample raw data (date, time, call-type, call duration) for each call activity in phone call log captured by the device.

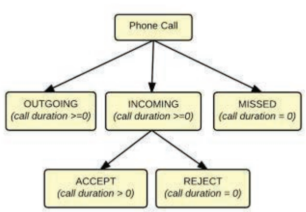

Figure 2: Overall classification of phone calls using call-type and call duration.

offers a unique opportunity to explore the potentiality of such information to model the call response behavior of individuals. In particular an individual's call response information (e.g., specific call time, call date, call type, call duration) in call logs provides raw contextual information about when she accepts, rejects or misses incoming calls. To an extent an individual's response to incoming calls is determined by the activities of an individual at the time of receiving the call. Many of these activities, as such call responses, follow a regular pattern at various times of the day and days of the week.

Let's consider a real life example of a mobile user, Peter, an executive officer. He regularly drives to work (an activity not recorded in the calendar) from 7:55AM to 8:40AM on Weekdays (Monday to Friday). The user rejects the incoming call in that time period. At other times, he may typically accept calls. Since the activity schedule and the corresponding behavior varies user to user over time-of-week, so the key challenge is to identify the like behavioral time segments that are used to producing call response rules of individuals utilizing phone call log data.

The temporal patterns of the call responses may provide the basis for modeling individuals call response behavior. In this paper, we investigate the potentiality of the phone call log information, particularly temporal information, to modeling individual's call response behavior.

A number of works have been done on call response rule-based interruption management system. Khalil et al. [1] use calendar information to infer user's activity and to automatically configure cell phones accordingly to manage interruptions. Seo et al. [2] proposed a context-aware configuration manager for smartphones PYP (Personalize Your Phone) to block a phone call without bothering the user. Kabir et al. [3, 9] proposed a novel rule-based technique that incorporates social context to minimize call interruptions. An intelligent context aware interruption management system has designed by Zulkernain [4].

The main drawback of the above approaches is that the rules used by the applications are not automatically discovered rather users need to define and maintain the rules manually by themselves. Moreover, users may not have the time, inclination, expertise or interest to maintain such rules. In this work, we aim to automatically discover individual's call response behavioral rules for developing an intelligent call interruptions management system. We envision that calling activity records in phone call logs are a rich resource to discover such rules automatically. Towards this goal, in this paper, we present our behavior modeling approach for generating individual's call response behavioral rules from phone call log.

Behavior Modeling

Log Pre-processing: The mobile phone records three types of call actions (INCOMING, MISSED and OUTGOING) in its call log. As outgoing call is not related to phone call interruptions, we ignore the outgoing call data. Figure 1 shows a snippet of raw log data except outgoing call captured by the mobile device. As can be seen, the user's actions in ACCEPTing and REJECTing calls are not directly distinguishable in INCOMING calls in the log. As such, we derive ACCEPT and REJECT calls by using the call duration. If the call duration is greater than 0 then the call has been

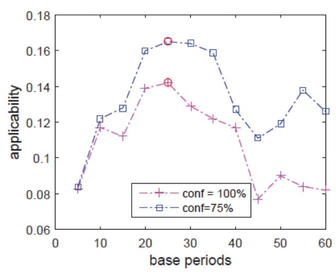

Figure 3: Optimal base period (minutes) selection for a sample user. The base period that produces the maximal (peak) applicability (marked by circle) for a particular confidence threshold is the optimal base period for that confidence level.

ACCEPTED; if it is equal to 0 then the call has been REJECTED. Figure 2 shows the overall classification of phone calls.

Temporal Analysis: As can be seen in Figure 1, phone log contains specific timestamp for each call action. However, human understanding of time is not precise, unlike digital systems. Our daily activities occur in time intervals rather than an exact time. For example, a person does not arrive at work or eat lunch exactly at the same time every day. There is always a time interval for routine behaviors, even it is a small interval, e.g., five minutes meeting. This temporal aspect needs to be considered in capturing the behavior patterns of individuals.

To take into account such temporal aspect, we follow the bottom-up approach in our model. First, each day of the week is divided into relatively small time slices (called the base period) say 5 minutes. Then we identify the dominant call response behavior for each time slice using a dominant threshold. This dominant threshold (rule confidence level) varies from user to user according to their preference on how interventionist they want the call-handling agent to be.

Once the dominant behavior has been identified for each time slice, call instances from adjacent time slices of like dominant behavior are aggregated dynamically and the corresponding time slices are merged to get the like behavior in a single longest possible time segment that is used to discover the call response rules. As the users daily schedule differ from day to day, in addition to the time of a day, we also take into account the specific day of the week to get pertinent rules.

The most appropriate segmentation will depend on the particular pattern of the user's activities. As we do not assume any prior knowledge of these activities, we therefore generate a number of sets of rules by iteratively varying the base time period. For each set of rules we compute the applicability which is the product of aggregate *support* and aggregate *time coverage*, where aggregate support is the fraction of the summation of the support count (the number of call instances) of all rules that satisfy the minimum confidence threshold among the maximum possible support considered, and the aggregate time coverage is the proportion of the time cover by those rules.

The base period that yields maximum applicability establishes the *optimal base period* for capturing best behavioral pattern of an individual. Figure 3 shows an example to identify the optimal base period. This base period may vary from user to user. The reason is that different users may have different activity patterns and thus different phone call response behaviors. The set of rules based on the *optimal time segmentation* represents individual user’s behavior.

Discovering Rules: Rules can be used to represent the call response behavior model. To discover such call response rules of an individual user using the produced optimal time segments, we employ the well-known association rule learning algorithm Apriori [6]. The reason is that other classification rule learners (e.g., decision tree) are ill-defined and non-deterministic in the sense that in general, one cannot guarantee the high predictive accuracy of the discovered classification rules [8]. In contrast, association-rule learning is well-defined and deterministic as it can discover precisely the rule set having support and confidence greater than

Table 1: Precision and Recall for high (max) confidence threshold (user X)

Confidence	100%
Precision	1
Recall	0.54

Table 2: Precision and Recall for comparatively low confidence threshold (user X)

Confidence	70%	60%
Precision	0.97	0.94
Recall	0.57	0.70

a user-specified threshold. According to the definition, *support* measures the frequency of association, i.e., how many times the specific type of call has occurred in a dataset, and *confidence* measures the strength of the discovered rules. As association rule learning allows a user to configure the confidence threshold value for creating rules, it reflects individual's preferences to modeling call response behavior. For instance, an individual may want the automated response handling agent to reject calls where in the past he/she had rejected calls more than, say, 70% of the time - that is, confidence threshold of 70%. Another individual, on the other hand, may want the agent to intervene if he/she had rejected calls in, say, 90% of past instances. This confidence threshold for creating rules will differ according to an individual's preference on how interventionist she wants the call-handling agent to be.

Experimental Evaluation

Data Collection. We have used two real datasets that consist phone call records of different mobile users. The first one is called 'Swin' dataset that was collected from individual mobile users by us. To do this, we have developed an Android mobile app which can collect the user's real current call log data on their mobile phone. Using this app we have collected call log data of 22 individual mobile users with various professions such as undergraduate students, post graduate students, university lecturers and industry professionals from August 2014 to September 2015. The second one is 'Nodobo' [7], a public release of mobile phone usage data that includes call records of 27 participants over a 5-month period. We have run our technique for all the 49 users of the above two datasets. As our approach is individualized, in this paper we have reported the experimental results of a sample user X (randomly selected).

Behavioral Rules. Below are some temporal rules for user 'X' generated by our behavioral modeling approach utilizing his phone call log as a context source. The rules are generated for the identified optimal base period 25 minutes with a minimum confidence threshold of 85%.

$$R1: DayTime \rightarrow Friday[16.15 \sim 17.30] \Rightarrow behavior \rightarrow reject$$
$$(Confidence = 100\%)$$
$$R2: DayTime \rightarrow Tuesday[18.20 \sim 20.00] \Rightarrow behavior \rightarrow missed$$
$$(Confidence = 88\%)$$
$$R3: DayTime \rightarrow Saturday[10.00 \sim 13.30] \Rightarrow behavior \rightarrow accept$$
$$(Confidence = 100\%)$$

Rule R1 states that user X always (100%) rejects calls between 16:15 and 17:30 on Fridays. Similarly, rule R2 states that the user misses most of the calls (88%) between 18:20 and 20:00 on Tuesdays. Finally rule R3 states that the user always (100%) accepts calls between 10:00 and 13:30 on Saturdays.

Accuracy Measurement: To evaluate the accuracy and effectiveness of our behaviour modelling approach, we use precision and recall as evaluation measures. Here precision calculates the fraction of used support instances in rules that are relevant, while recall calculates the fraction of relevant instances that are captured by produced output rules. Table 1 and Table 2 report the precision and recall with varied confidence. The results show that precision decreases with the decrease of confidence threshold as per definition. In contrast, recall increases with the decrease of confidence threshold because lower threshold value satisfies more rules for output. Thus the total support in rules increases and as a result recall increases. The results also show that our modeling approach discovers behavioral rules of an individual with high accuracy (in terms of precision and recall) for a particular confidence threshold.

Conclusion & Future Work

In this paper, we have explored the potentiality of phone call log as a context source for modeling call response behavior of individual mobile users. We have presented our initial work that utilizes the contextual information of phone call log to generate call response behavioral rules i.e., when a user accepts, rejects or misses an incoming call. In particular, our approach incorporates temporal and individual aspects to model call response behavior. Our behavioral model can assist developers to building automated behavior-oriented call interruption management system for the benefit of end users. Our preliminary experimental results on real datasets demonstrate that context information in phone call log has great potential to model individual's phone call response behavior with high accuracy. In future work, we plan to study and incorporate additional contexts such as social relationships between the caller and receiver to further enhance our call response behavior model.

References

1. A Khalil, K Connelly. 2005. Improving Cell Phone Awareness by Using Calendar Information. In Human-Computer Interaction-INTERACT, Springer, 588-600.
2. Sin-seok Seo, Arum Kwon, Joon-Myung Kang, John Strassner, and James Won-Ki Hong, 2011. PYP Design and Implementation of a Context-Aware Configuration Manager for Smartphones, Proceedings of the 1st International Workshop on Smart Mobile Applications (SmartApps' 11).
3. Muhammad Ashad Kabir, Jun Han, Jian Yu, and Alan Colman. 2014. User-centric social context information management: an ontology-based approach and platform. Personal Ubiquitous Computing. 18, 5 (June 2014), 1061-1083.
4. Sina Zulkernain, Praveen Madiraju, Sheikh Iqbal Ahamed and Karl Stamm. 2010. A Mobile Intelligent Interruption Management System. Journal of Universal Computer Science (JUCS), 16, 15: 2060—2080.
5. Vijay Srinivasan, Saeed Moghaddam, Abhishek Mukherji, Kiran K Rachuri, Chenren Xu and Emmanuel Munguia Tapia. 2014. MobileMiner: Mining Your Frequent Patterns on Your Phone, Proceedings of the ACM International Joint Conference on Pervasive and Ubiquitous Computing, 389-400.
6. Rakesh Agrawal and Ramakrishnan Srikant, 1994. Fast Algorithms for Mining Association Rules, Proceedings of the 20th International Conference on Very Large Databases (VLDB), 487-499.
7. Stephen Bell, Alisdair McDiarmid, and JamesBell Irvine. 2011. Nodobo: Mobile phone as a software sensor for social network research. In Vehicular Technology Conference (VTC Spring), 1-5.
8. Alex A Freitas. 2000. Understanding the crucial differences between classification and discovery of association rules: a position paper. ACM SIGKDD Explorations Newsletter, 2(1):65–69.
9. M. A. Kabir, J. Han, J. Yu and A. Colman, 2015. Inferring User Situations from Interaction Events in Social Media. The Computer Journal, Section C: Computational Intelligence, Machine Learning and Data Analytics, Oxford University Press, Oxford, UK, 58(9).