# Understanding Recency-Based Behavior Model for Individual Mobile Phone Users

**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'17 Adjunct*, September 11-15, 2017, Maui, HI, USA

ACM ISBN 978-1-4503-5190-4/17/09.
https://doi.org/10.1145/3123024.3124570

## Abstract
Mobile phone log data is not static as it is progressively added to day-by-day according to individual's behavior. The goal of this position paper is to highlight the issues of *traditional behavior modeling* utilizing phone log data and to describe the key aspects that constitute the foundation of our *recency-based behavior modeling* for individual mobile phone users to overcome such issues.

## Author Keywords
Mobile data mining; User behavior modeling; Incremental rule mining; Recency;

## ACM Classification Keywords
H.2.8 [Database Applications)]: Data mining; H.3.4 [Systems and Software)]: User profiles and alert services

## Introduction
In this position paper, we (1) highlight the issues of traditional behavior modeling, and (2) introduce a concept of 'recency-based behavior modeling'. Our approach not only takes into account the *recency-based rules* (rules based on recent behavioral patterns) but also removes the *outdated rules* according to individual's daily mobile phone usages behavior. In this paper, we discuss the key aspects that constitute the foundation of our *recency-based behavior modeling* for individual mobile phone users.

*Behavior Modeling Issues*
Mobile phones have become increasingly ubiquitous and powerful. Now-a-days, almost 95% people in the world use mobile phones [8] and the phones are, for most of the day, with their owners as they go through their daily routines [9]. Modeling an individual's behavior utilizing phone log data may assist them in their daily activities through capabilities, such as mobile notification management [7], phone call management [10, 12], application recommender systems [14]. A key step of such modeling is to extract *behavioral rules* of individuals' based on related contexts such as temporal, location etc.

Currently, researchers use *static length* of log data, e.g., 12 months, for discovering rules in order to model mobile phone usage behavior [7, 10, 14]. The problem is that the produced rules utilizing the static length of log data do not guarantee the present behavior of a user, as an individual's behavior changes over time. Let's consider last 12 months call log data of a mobile phone user. Assume that as per log data the user has a call 'reject' behavioral pattern on Monday[10:00AM-12:00PM] as she used to have a regular meeting at that time. Recently she has no meeting at that time period on Monday and she typically 'accepts' incoming phone calls. So for this example, the past 'reject' behavioral pattern, even with high evidence according to log data, is not meaningful to predict her future behavior. Therefore, producing and managing updated rules according to individual's *recent behavioral patterns* is a key issue for individual's behavior modeling in the real world applications.

To produce rules that express users' present behavior, a number of research [5] [10] have used the behavioral patterns of recent mobile phone log data to predict the future behavior rather than the patterns derived from the entire historical logs. For example, Lee et al. [5] have studied the mobile phone users' calling patterns and used last three months call logs. Phithakkitnukoon et al. [10] have presented a model for predicting incoming and outgoing calls and assumed latest 60 days call logs data to model future call activities. However, such arbitrary data length is also *static* and is not meaningful to predict future mobile phone usages behavior as the time frame of the data length depends on when in the recent past behavior of an user has been changed significantly. Even the behavior of an individual user varies from day-to-day as the daily schedules are not identical in the real world. For instance, a user has a meeting on every Friday [2:00PM-3:00PM] and rejects the incoming calls during that time, but on other days he is free at that time and typically accepts the incoming calls. But if we do not differentiate between days-of-the-week, the other days' unlike activities will mask the dominant behavior on Friday, and we would thereby miss the Friday's reject rule. Therefore, it is very difficult to assume a length of log data *statically* for modeling individual's mobile phone usages behavior.

## Key Aspects of Recency-Based Behavior Model
To address the above mentioned issue in behavior modeling, we introduce a concept of 'recency-based behavior modeling' that takes into account not only the recency-based rules but also removes the outdated rules according to individual's daily mobile phone usages behavior. In our approach, we dynamically identify data length to produce rules that capture a user recent behavior patterns. We further merge these rules with the existing rule set that is produced from initial log data. Hence, the key aspects that constitute the foundation of our *recency-based behavior modeling* are as follows:

*Incremental Mining*
Incremental mining is the basis of our recency-based behavior modeling. As the amount of mobile phone log data increases day-by-day according to individual's behavior, the incremental mining and updating the rules for such a dynamic dataset is important for modeling mobile phone usages behavior of individuals. The updates may not only invalidate some existing rules but also make other rules relevant. Several incremental mining techniques have been proposed for mining rules in a dynamic dataset. For instance, frequent pattern based [2, 3, 15, 17], three-way decision based [6, 18], and probability-based [1, 16] techniques efficiently maintain association rules of a dynamic dataset. These techniques produce updated rules by taking into account the incremental dataset and the knowledge of existing rules produced from initial dataset. However, such incremental techniques do not take into account the *freshness of rules* (e.g., recency) that represents the present behavior of an individual.

*Recent Behavioral Pattern Based Rule Generation*
A behavioral rule is represented as $X \rightarrow Y$, where X is defined as the antecedent and Y as the consequent. The antecedent contains contextual information and the consequent contains individual's mobile phone usages behavior. This means that rules can be in the form $X \rightarrow Y$ but not in the form of $Y \rightarrow X$. To produce such rules based on recency, we take into account most interesting behavioral patterns of individuals according to related contexts.

Since human behavior changes over time, the most *recent pattern* is more interesting and significant than older ones for predicting individual's mobile phone usages behavior. According to [11], the recent trend of the caller/user's calling pattern has higher correlation to the future pattern than the pattern derived from the entire historical log data. Therefore, identifying a particular length of log data that contains recent behavioral patterns is the key for producing recency-based rules of individual mobile phone users. For this, if we take into account only a short length (e.g., last week's data) as indicative of recent behavior, there may not be enough data instances in that time period to infer a valid rule. Creating rules based on observations with so little "support" is unlikely to be effective. On the other hand, if we take into account comparatively longer lengths (e.g., last 6 months data) as indicative of recent behavior, we could get greater *support* but it might result a greater *behavioral variations* thus decrease the confidence of some expected rules. As a consequence we may miss these rules because of not satisfying the confidence threshold set by individuals. Therefore, the *optimal length* of recent log data that reflects the present behavior of individuals needs to be identified for producing recency-based rules. As we are not interested to assume such length statically, the optimal length can be identified by *measuring behavioral similarity* between adjacent weeks started from the most recent week by taking into account the behavioral patterns in different contexts for each day-of-the-week. We take into account users' behaviors on a weekly basis, as time-of-the-week is an important aspect of user behavior in a mobile-Internet portal [4]. Figure 1 shows an example for changing behavioral patterns for each day-of-the-week.

According to Figure 1, week $n$ is the most recent week and week $s$ is the boundary of recent behavioral patterns, that is, the behavioral patterns based on related contexts before week $s$ are considered as past behavior and the behavioral patterns after week $s$ up to week $n$ ([week $s$ - week $n$]) are considered as present behavior of users. If there is no change in behavioral patterns from week 1 (beginning of log data) to week $n$, then the behavioral patterns for that particular days-of-the-week (e.g., Thursday and Saturday

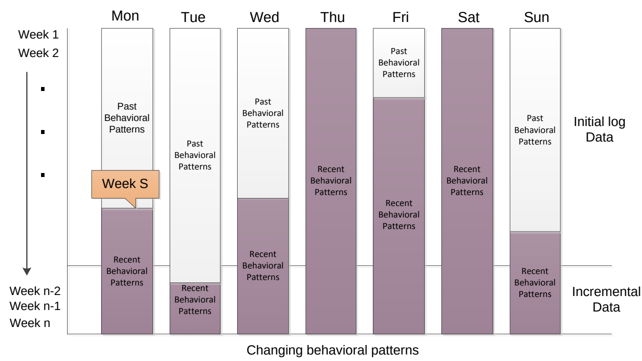


**Figure 1:** Changing behavioral patterns for each day-of-the-week

patterns in Figure 1) are considered as recent patterns. We utilize such variable length of recent log data for producing recency-based rules. It varies from user-to-user depending on how the user's behavior changes over time-of-the-week in different contexts.

Once the recent log data [week $s$ - week $n$] has been identified, we produce rules utilizing this data. To produce behavioral rules, we employ the rule learning technique [13] that produces effective non-redundant behavioral rules in multi-dimensional contexts consisting of both general and specific rules (exceptions of general rules). The rule learning technique allows a user to configure the confidence threshold value for creating rules that differ according to an individual's preference on how interventionist she wants the agent to be. For example, one person may want the agent to reject calls where in the past he/she has rejected calls more than, say, 80% of the time - that is, at a threshold of 80%. Another individual, on the other hand, may only want the agent to intervene if he/she has rejected calls in, say, 95% of past instances

*Updating Rule-set based on Recency*

This is the final step of our recency-based behavior model. In our approach, once we have produced recency-based rules utilizing a dynamic length of recent log data, we merge these rules with existing rule-set (produced from initial log data) to output a *complete set of updated rules* for each individual. To do this, we first identify the outdated rules from existing rule-set using the consequent of recency-based rules. If a rule $X \rightarrow Y$ has the similar $X$ but different $Y$ in both rule-sets, then the rule in existing rule-set has been identified as an 'outdated rule' that does not represent the present behavior of an individual. To get a complete set of updated rules, we first remove all the outdated rules from the existing rule-set and add all the recency-based rules. This complete updated rules-set not only contains all the useful rules of an individual mobile phone user from week $1$ to week $n$ but also expresses recent behavioral patterns that will be helpful for modeling mobile phone usages behavior in the real world applications.

## Discussion

To the best of our knowledge, this is the first recency-based behavioral rule extracting study for modeling mobile phone usages behavior. In this paper, we have discussed the key aspects that constitute the foundation of our *recency-based behavior modeling* for individual mobile phone users utilizing their phone log data. We believe that our concept of recency-based behavior modeling helps both the researchers and application developers for predicting behavior of end mobile phone users according to their needs in various real-world applications, such as call interruption handling, notification management systems, recommender systems etc. However, it has some minor limitations. Firstly, we take into account users' behaviors on a weekly basis for identifying recent behavioral patterns of individuals, as the behavior is influenced by time-of-the-week [4]. However, in some cases, monthly or yearly basis behavioral patterns might be meaningful. Secondly, though we are able to produce recency-based rules of individuals by identifying an optimal length of recent log data, we need existing knowledge (rules-set) to get a complete set of updated rules for the whole log period.

## Conclusion

In this paper, we have highlighted the issues of traditional behavior modeling utilizing mobile phone log data and have described the key aspects of recency-based behavior modeling for individual mobile phone users. We believe that our discussion opens a promising path for future research on modeling user behavior based on their daily work routine. Although we choose the phone call behavior as an example, it is also applicable to other application domains.

## REFERENCES


1. Ratchadaporn Amornchewin and Worapoj Kreesuradej. 2009. Mining Dynamic Databases using Probability-Based Incremental Association Rule Discovery Algorithm. *J. UCS* 15, 12 (2009), 2409–2428.

2. David W Cheung, Jiawei Han, Vincent T Ng, and CY Wong. 1996. Maintenance of discovered association rules in large databases: An incremental updating technique. In *Data Engineering, 1996. Proceedings of the Twelfth International Conference on*. IEEE, 106–114.

3. David Wai-Lok Cheung, Sau Dan Lee, Ben Kao, and others. 1997. A General Incremental Technique for Maintaining Discovered Association Rules.. In *DASFAA*, Vol. 6. 185–194.

4. Martin Halvey, Mark T Keane, and Barry Smyth. 2006. Time based patterns in mobile-internet surfing. In *Proceedings of the SIGCHI Conference on Human Factors in computing systems, Montreal, Quebec, Canada, 22-27 April, pp.∼31–34. ACM, New York, USA*.

5. Seunghwan Lee, Jungsuk Seo, and Geehyuk Lee. 2010. An adaptive speed-call list algorithm and its evaluation with ESM. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 2019–2022.

6. Yao Li, Zhi-Heng Zhang, Wen-Bin Chen, and Fan Min. 2015. TDUP: an approach to incremental mining of frequent itemsets with three-way-decision pattern updating. *International Journal of Machine Learning and Cybernetics* (2015), 1–13.

7. Abhinav Mehrotra, Robert Hendley, and Mirco Musolesi. 2016. PrefMiner: mining user's preferences for intelligent mobile notification management. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, 1223–1234.

8. Mert Ozer, Ilkcan Keles, Hakki Toroslu, Pinar Karagoz, and Hasan Davulcu. 2016. Predicting the Location and Time of Mobile Phone Users by Using Sequential Pattern Mining Techniques. *Comput. J.* 59, 6 (2016), 908–922.

9. Veljko Pejovic and Mirco Musolesi. 2014. InterruptMe: designing intelligent prompting mechanisms for pervasive applications. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, 897–908.

10. Santi Phithakkitnukoon, Ram Dantu, Rob Claxton, and Nathan Eagle. 2011. Behavior-based adaptive call predictor. *ACM Transactions on Autonomous and Adaptive Systems (TAAS)* 6, 3 (2011), 21.

11. Santi Phithakkitnukoon and Carlo Ratti. 2010. A recent-pattern biased dimension-reduction framework for time series data. *Journal of Advances in Information Technology* 1, 4 (2010), 168180.
12. Iqbal H Sarker, Alan Colman, Muhammad Ashad Kabir, and Jun Han. 2016. Phone call log as a context source to modeling individual user behavior. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct*. ACM, 630–634.
13. Iqbal H. Sarker, Ashad Kabir, Alan Colman, and Jun Han. 2017. An Approach to Modeling Call Response Behavior on Mobile Phones Based on Multi-dimensional Contexts. In *Proc. of 2017 Int. Conf. on Mobile Software Engineering and Systems (MOBILESoft), Argentina.* 91–95.
14. Vijay Srinivasan, Saeed Moghaddam, Abhishek Mukherji, Kiran K Rachuri, Chenren Xu, and Emmanuel Munguia Tapia. 2014. Mobileminer: Mining your frequent patterns on your phone. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, 389–400.
15. Shiby Thomas, Sreenath Bodagala, Khaled Alsabti, and Sanjay Ranka. 1997. An Efficient Algorithm for the Incremental Updation of Association Rules in Large Databases.. In *KDD*. 263–266.
16. Panita Thusaranon and Worapoj Kreesuradej. 2015. A probability-based incremental association rule discovery algorithm for record insertion and deletion. *Artificial Life and Robotics* 20, 2 (2015), 115–123.
17. Baowen Xu, Tong Yi, Fangjun Wu, and Zhenqiang Chen. 2002. An incremental updating algorithm for mining association rules. *Journal of Electronics (China)* 19, 4 (2002), 403–407.
18. Zhiheng Zhang, Yao Li, Wenbin Chen, and Fan Min. 2014. A three-way decision approach to incremental frequent itemsets mining. *JOURNAL OF INFORMATION &COMPUTATIONAL SCIENCE* 11, 10 (2014), 3399–3410.