Understanding Mobile Banking Usage Behavior: An Analytics Approach

Luvai F. Motiwalla  
*University of Massachusetts - Lowell*, luvai_motiwalla@uml.edu

Mousa Albashrawi  
*University of Massachusetts - Lowell*, mousa_albashrawi@student.uml.edu

Hasan Kartal  
*University of Massachusetts Lowell*, hasan_kartal@student.uml.edu

Follow this and additional works at: [http://aisel.aisnet.org/sigdsa2016](http://aisel.aisnet.org/sigdsa2016)

**Recommended Citation**  
[http://aisel.aisnet.org/sigdsa2016/14](http://aisel.aisnet.org/sigdsa2016/14)

This material is brought to you by the Special Interest Group on Decision Support and Analytics (SIGDSA) at AIS Electronic Library (AISeL). It has been accepted for inclusion in Proceedings of the 2016 Pre-ICIS SIGDSA/IFIP WG8.3 Symposium: Innovations in Data Analytics by an authorized administrator of AIS Electronic Library (AISeL). For more information, please contact elibrary@aisnet.org.
Understanding Mobile Banking Usage Behavior: An Analytics Approach

Research-in-Progress

Luvai F. Motiwalla 
UMass Lowell 
OIS Department 
1 University Ave, Lowell, MA 
luvai_motiwalla@uml.edu

Mousa Albashrawi 
UMass Lowell 
OIS Department 
1 University Ave, Lowell, MA 
mousa_albashrawi@student.uml.edu

Hasan B. Kartal 
UMass Lowell 
OIS Department 
1 University Ave, Lowell, MA 
hasan_kartal@student.uml.edu

Abstract

Understanding mobile banking (MB) system usage is of great value and significance to both academicians and practitioners. MB use has increased to 30% of all bank users, generating over $500 m. annual revenue for bank according to WSJ. This paper provides a novel approach to study MB usage by combining data analytics technique of cluster analysis with a field survey. These users were segmented into homogeneous groups by using cluster analysis with their banking activity. In our study, we use cluster analysis to segment 4478 users based on their usage behavior from their log file activities captured with a MB system deployed by a mid-size bank in U.S. After clustering the banking transactions on a variety of MB usage attributes such as number and types of activities, we have created three homogeneous user groups who will be surveyed to determine the MB system success with IS success model. This field study will also gather their demographic background and experience with the MB system. The survey results will be analyzed while retaining the homogeneous groups allowing inter-group comparisons on the constructs from the IS success model. Theoretical contributions from this study are the innovative use of data analytics for segmenting user sample on the usage variable before the behavioral study. This reduces the risk of sample heterogeneity bias, a priori, before the field study. This approach will provide a deeper understanding of MB usage than prior studies that have used heterogeneous samples and help banks better decisions on MB system features that are more attuned to the needs of the users.

Keywords: IT system usage, mobile banking, clustering, analytics

Introduction

System usage has been studied extensively in IS literature (Burton-Jones & Straub 2006). One limitation is that more studies have used subjective than objective measures. While a subjective measure like intention to use the system is important, it is not as powerful as an objective measure like the actual use of the system (Eckhardt et al. 2013). Actual use measures frequency of use (Lu and Gustafson, 1994), feature diversity (Brynjolfsson et al. 2011), and feature utility (Head & Ziolowski, 2012) are best observed through the objective measures. Actual use captures user behavior through the system log files (Walldén et al. 2015) or through newer psychophysiological tools like eye-tracking (Eckhardt et al. 2013). Another limitation of
system usage studies is the use of a single heterogeneous sample (Schacht et al. 2015). Many adoption studies have self-selection and self-reporting biases that limit the validity of IS adoption studies (Lee et al. 2003). When samples are segmented they are mostly segmented on user demographic variables like age or gender or education (Schacht et al. 2015), or geographic region (Mallat et. al. 2006), or system use experience (Sell et al. 2012) used either as control variables or finite mixture partial least squares (FIMIX-PLS) approach. We have not found a study that has segmented the users, using data analytic techniques, based on their usage activity from the system logs.

Our study focuses on the mobile banking (MB) system usage, after initial adoption. Mobile technology adoption studies have focused mostly on the intention to use with either the technology acceptance model (TAM) or the unified theory of acceptance and usage of technology (UTAUT), derived from IT adoption theories (Gao & Bai, 2014). Few studies have focused on mobile system success, especially with objective measures (Eckhardt et al. 2013). Also, MB usage has grown tremendously in last decade, and banks see more value from MB post-adoption usage rather than adoption (Huang 2014). MB is considered a strategic technology by the banking industry for cost savings, customer engagement, retention and satisfaction (FiServ Report, 2016). Customer acceptance of MB system has been growing fast but varied; some users have accepted it quickly and are even willing to pay for it, while others are just getting accustomed to this technology (Crowe et al. 2015). The one-size-fits-all approach is not feasible today when companies are deploying new strategies to fit mobile technologies to the different needs and preferences of their users. Overloading systems with functionality can lead to feature fatigue and under usage (Head & Ziolowski, 2012). Therefore, understanding usage with segmented groups (Head & Ziolowski, 2012) is similarly critical for the banking industry.

MB has become an essential component of everyday banking and has shown to increase customer loyalty with an average of 2.3 product holdings per MB customer vs. 1.3 product holdings per non-MB customer (FiServ Report, 2016). The same study found that MB users account for 14.4 percent of the customer base in the banks, they drive more than 39 percent of the total revenue. Statista¹, a statistics portal, reports MB usage has increased from 35 million users in 2009 to 89 million users in 2014 and is projecting an increase to 111 million users by the end of 2016. This is a 217% increase in MB users over a period of seven years. MB usage, based on an analysis from two banking industry studies (Crowe et al., 2015; FiServ Report, 2016), can be categorized from lean to rich activities. Lean activities like checking account balances (94%), downloading MB Apps (71%), the number of transfers between accounts (61%) form a bulk of MB usage activity but do not generate much revenue for the banks (Crowe et al. 2015). Rich activities, on the other hand, such as bill payment amounts (48%), money transfer amounts between bank users (25%), and mobile deposit amounts (48%) that generate on average 5-6 times higher fee revenues and reduce customer servicing costs ($1.5B); MB costs only a few cents per transaction (FiServ Report, 2016). In addition, most rich MB activities involve the use of multiple interactions between users and the system and a higher level of cognitive tasks. This indicates that the analysis of MB usage patterns is critical.

A major contribution of this paper is combining the objective (data analytics) with the subjective (survey) measures for a better understanding of MB success from users’ actual usage. Cluster analysis is used in phase one to segment users into homogeneous groups and followed by survey with the same users, using the IS success model (DeLone & McLean, 2003), in phase two. Extant mobile usage studies have relied purely on subjective measures using TAM or UTAUT models which focus on the antecedents of behavioral intention to use during initial adoption or continued use intention (Gao & Bai, 2014). The IS success model was chosen for its emphasis on system success through the actual use and user satisfaction influenced by the quality of the system, information, and service. The inclusion of satisfaction and use as dependent variables makes the IS success model appropriate for our study. Further, no prior studies have used a dependent variable (usage) from the IS success model for segmenting the user sample. Segmentation of users allows us to compare the MB systems’ success factors, satisfaction, and use among the user groups. In sum, our behavioral-analytics approach provides a more comprehensive and insightful analysis of MB usage and success which, to the best of our knowledge, has not been used by any prior system usage studies. The rest of the paper is organized as follows. After a review of related work in the next section, we discuss our research model, research approach and preliminary analysis from phase one analytics study and our plans.

for the phase two study. We conclude with the potential theoretical and practical contributions of our research.

Related Work

One key goal of this study is to evaluate MB usage by segmenting users into homogenous groups from their actual usage behavior. Different consumer segments may value different success factors (Head & Ziolkowski, 2012; Wind 1978) with their MB use. Heavy users may be utilizing the system for different reasons than light users. Recent studies have reported that lack of user segmentation has introduced statistical validity threats to empirical testing of behavioral study models due to heterogeneity in the samples which bias the study results that lead to invalid conclusions (Becker et al. 2013; Schacht et al. 2015). Both these studies have proposed grouping users after data collection from surveys by using advanced statistical techniques like prediction-oriented segmentation with partial least squares (PLS-POS) or finite mixture partial least squares (FIMIX-PLS) approach with IT adoption models to detect unobserved segments and have found that segmentation reduces bias in parameter estimation and inferential errors. These studies suggest user segmentation in a behavioral study is important because it enables researchers to better understand the factors influencing usage (Becker et al. 2013), improve MB usage by considering individual needs (Schacht et al. 2015), and increase the mobile system usage (Head & Ziolkowski, 2012). In this study, we are therefore segmenting users into apriori groups based on their MB system usage with cluster analysis followed by a field study using the IS success model.

Cluster analysis has been a popular method of segmenting customers in marketing research (Wind, 1978) and has been used in IS research successfully for discovering user behaviors in online shopping based on website visits, clicks or orders (Song et al. 2001) as well as studying switching behaviors on website portals (Lee et al. 2003) and more recently understanding customer migration behavior in mobile service usage and revenue patterns (Bose & Chen, 2015).

Clustering users on their mobile usage patterns, before determining the antecedents or consequents affecting system usage, can provide a better understanding of system success factors than cross-sectional studies which assume models can be generalized to user population (Wu et al. 2015). A “one-size fits all system usage model can be misleading as different consumer segments (segmented by perceived feature utilities rather than demographics) may value different constructs or experience different causal relationships in the model” (Head & Ziolowski, 2012). User homogeneity provides a better understanding of the needs of different user groups and their usage of the system functions (Walldén et al. 2015). For example, heavy users will have different expectations and perceptions of the same system as compared to the light users. Heavy users use the system frequently and access basic to advance features of the system, and have a different level of confidence and system use experience (Becker et al. 2013). This often leads to a different level of satisfaction and use, as compared to the light users. Similarly, it is possible that light users who are interested in limited MB features they use, leading to the same level of satisfaction as a heavy user. Our research will address these issues with a more homogeneous user segments. Sample heterogeneity is a likely cause of validity problems due to Type I or Type II errors and affects IS adoption or usage study results (Becker et al. 2013). Incorporating segmentation methods from the marketing domain (Wind, 1978) provides a more comprehensive and richer understanding of mobile user preferences and motivations for use (Head & Ziolowski, 2012). Similarly, Schacht et al. (2015) observe that IS adoption studies have used segmentation from marketing research to group either by users’ system experience, age or by their adoption stage (Rogers, 1983). Segmenting a large diversified user group is important when system usage is voluntary, as in mobile banking. Unlike employees, consumers are not mandated to use these systems (DeLone & McLean, 2003). Determination of system success is more difficult when users’ motivations and reasons for use are diverse.

Burton-Jones and Straub (2006) have developed a measurement scale which categorizes system usage from lean to rich. Lean measures determine usage based on a simple measure like use/nonuse, the frequency of use of a system features, and duration of use. While rich measures determine usage based on the extent to which a user understands and employs the system features to accomplish a task. They reflect on the users’ cognitive absorption of system features and use of the system for advanced tasks that involve multiple interactions between system and users to accomplish a task. We have adapted their definition of lean and rich measures to categorize the MB activities and use them to cluster our user sample into the light, moderate and heavy segments. The amounts of money transfers, mobile deposits, and bill payments were
defined as rich measures because they exhibit a higher level of users’ cognitive absorption with the system features and generate more value for banks. The frequency of logins, number of balance checks, and deposits were defined as lean measures because they focus only on the presence and extent of the MB system use.

Cluster analysis has been used to identify consumer segments on the basis of product usage variables in marketing research (Schaninger et al. 1980). A priori clustering based on their usage behavior first allows us to better understand MB users’ behaviors by identifying smaller user segments without violating their individual privacy. Second, it allows the banks to devise certain MB features, services, and incentives that are customized to the needs of these groups for better success outcomes. Finally, it allows us to test the IS success model factors with a more homogeneous user group, thereby reducing the chance of statistical validity errors (Type-I & Type-II) often caused by observed or unobserved heterogeneity (Becker et al. 2013). This a priori segmentation with cluster analysis into groups based on their MB usage patterns will provide a more comprehensive and richer understanding of user perceptions on MB system, information and service qualities and its’ influence on use and satisfaction with the IS success model.

IS Success Model

IS Success model (Delone & McLean, 2003) has been used by over 300+ evaluation studies to determine the success of information systems with users. It determines success by observing the influence of the three quality factors--system, information, and service quality--on system use, user satisfaction, and individual and organizational benefits (tangible and intangible) derived from using the systems.

The IS success model has been used in many different system usage studies and has been popular in mobile usage studies. It has been used for investigating success factors of mobile systems in healthcare (Chatterjee et al. 2009), studying the impact of trust and satisfaction in mobile banking (Lee & Chung, 2009), and understanding user satisfaction, trust, flow and continued usage intention with mobile payment systems (Zhou, 2013). These studies have generally found that systems quality which measures technical success, information quality which measures semantic success and service quality which measures customer service provided by mobile systems affects user satisfaction positively and increases use of the system. This suggests that the IS success model has validity in measuring the success of mobile systems and hence can be used in the context of our study. However, the majority of these studies have used it with a single non-segmented user sample. There is no analysis available on how the IS success factors affect different user segments. For example, each user segment may be affected differently by the three quality factors of the model. A higher system quality may lead to higher user satisfaction in one segment and lower in another segment. This leads to three general research questions: 1) “What segments of mobile banking users exist and how data analytics characterizes these segments?” 2) “What is the demographic profile of each user segment?” and 3) “How are different user segments influenced by the factors of IS Success model? Our study, like Zhou (2013) study, does not evaluate the individual or organization benefits at this stage because MB usage is relatively new and most users may not have sufficient time to reflect on their benefits from this technology.

In this study, our focus is more on evaluating the use and satisfaction differences amongst the segments. Petter et. al. (2012) in their review of evaluation of IS success model state that while systems role in the organization has changed over its 60-year history in organizations, factors used in evaluating system success has not changed. System, information, service quality, use, user satisfaction, individual benefits and organization impacts are still relevant in today’s customer-focused system era. But, higher levels of interaction with the system and empowerment of users has increased the complexity of measuring system success and changed the metrics relevant to these factors. Further, the personalization of systems must recognize the varying perceptions of different user groups and individuals. One group may see a system as successful while another group may see it as a failure (Bartis and Mitev, 2008).

Research Model

Analyzing survey data by the user segments, as shown in Figure 1, will improve the validity of our results and allow comparisons of system success amongst the light, moderate and heavy user groups. To the best of our knowledge, no prior research has used this approach of segmenting users in the sample with the data analytics approach on the system usage construct (Burton-Jones & Straub, 2006) or with the IS success
Understanding Mobile Banking Usage Behavior

model (DeLone & McLean, 2003). We believe our approach will provide a better understanding of the MB system usage.

We plan to determine the success of MB system by comparing the factors with different user segments determined from the cluster analysis into heavy, moderate and light users by using the lean to rich measures of MB activities. Assuming our clustering reveals three categories, as shown in Figure 1, our hypothesis will recognize different perceptions of system success amongst the segments; for example, heavy user segment may be highly satisfied and use the MB system for its information and service qualities, while moderate user segment may be somewhat satisfied and use the MB system for its system quality. On the other hand, light users may not be satisfied with the system due to its system quality. Our study will attempt to reveal the differing influences of the IS success factors on the different groups and identify whether significant differences exist amongst the segments. In addition, the descriptive analysis will provide details on the user demographic profiles of each segment in terms of age, gender, education and work status. A more in-depth theory building and hypothesis will be developed in our completed research paper.

Research Method and Preliminary Analysis

Our research method is designed to encompass two data collection mechanisms: user log file data followed by a field survey. In phase one, we have collected the log data files of MB usage from the bank’s mobile billing analytics and reporting file. This file contains detailed MB usage data on 51 attributes such as customer identifier number, name, date of first registration, number of activities on various banking transactions like balance checks, transfers, bill payment, deposits to the total amounts of deposits, payments, transfers, and bill payments all of which capture the use of the MB system. Use, which is a dependent variable of the IS success model, is used to segment users with cluster analysis.

Cluster analysis is an unsupervised machine learning method for grouping objects that have similar characteristics based on their distance between objects from a mean vector. It groups a data matrix composed of observations (rows) and variables (columns) into homogeneous segments without any data transformation as in principal component analysis (Giudici & Figini, 2009). A variety of clustering techniques are available depending on the goals of clustering (Berkhin, 2006). Broadly, they are categorized into hierarchical and partitioning algorithms. The hierarchical technique either uses agglomerative or divisive algorithms where objects are divided into parent-child relationships using trees. While flexible and easy to handle, they are vague on the termination criteria and work on the presumption that clusters consist of similar objects. Partitioning algorithms use iterative learning heuristics like probabilistic, k-medoid and k-mean methods to fit objects into homogeneous clusters by revisiting clusters after every iteration formation, gradually improving the clusters.

In phase two, an online survey will be conducted by the bank with the same users on their experience with the system using independent variables from the IS success model to understand their perceptions of the quality, information, and service factors, as well as their use and satisfaction with the system. All constructs items have been adapted from previous research to ensure face validity. The items will be measured via a
7-point, Likert-scale with 7 “Strongly agree” and 1 “Strongly disagree”. Quality factors (system, information, and service) and satisfaction are adapted from Zhou (2013). Survey responses will be matched on users’ email address from their log data by the bank, anonymized to prevent identity disclosures and sent to us for data analysis. We will analyze survey data by each user cluster from phase one. This approach allows us to determine system usage success by segment and determine whether significant differences exist or not between the groups. It also allows us to minimize the threats of statistical validity from sample heterogeneity (Becker et al. 2013) providing a better understanding of MB usage.

### Data Preparation and Descriptive Statistics

We have analyzed eight months of observations from the MB system log file, up to the month when we survey the users. From this file we extracted the seven attributes which help us understand the usage behaviors of a user; the eight attribute was their registration number. We divide these attributes into two categories, lean and rich measures, which were adopted from the Burton-Jones and Straub (2006, pg. 232) paper. In our context, the lean measures indicate the extent of use like the number of times a user requests money transfers, bill payments, mobile deposits, or transactions like balance checking. The rich measures indicate the degree to which the system was used for banking tasks that generate value for the bank and exhibit confidence of user in the system. The amount of money processed in a transaction like money transfers, bill payments, mobile deposits reflect rich usage of the MB system. The initial dataset contained 5,116 unique registrants. However, some users were early registrants who have the full eight months of observations, while the other users are late registrants who have limited observations varying from one to eight months of usage activity. To improve reliability, we analyzed the data for registrants with a minimum of six-month activity going backward from the survey date. To have a fair representation of varying (six, seven and eight months) lengths of the accounts, we aggregated the log data and scaled to monthly usage, that is generated by the sum of the observed values for a user divided by the number of months. The final dataset was reduced to 4,478 registrants each with an observation corresponding a value of monthly average.

In an average month, many users have less than one activity in many attributes except in the number of other activities. The averages for the mobile transfer amounts is less than $500, bill payment is less than $90 and submitted deposit is just above $130. The standard deviations of each attribute are high, as many as three to five times mean values, indicating a wide usage diversity in our sample supporting the use of cluster analysis to group MB system users into more homogenous segments, discussed next.

### Phase One: Clustering Analysis

We used cluster analysis for segmenting users into homogenous groups based on their MB usage behaviors. From the 4,478 observations, we first removed 740 observations with zero MB activities from our dataset. The remaining 3,738 observations were normalized on a scale of 0 – 1 and clustered with R software producing a 2-cluster solution, 3-cluster solution and 4-cluster solution as observed with the “elbow criterion” method. However, a visual inspection of the dendrogram generated by Ward’s hierarchical clustering algorithm indicated the validity of three cluster solution in our dataset. This process is called a reciprocal two-stage procedure, and it ensures the clustering results are stable and consistent.

To ensure a more balanced and reliable clustering solution, we employed two clustering techniques; K-means and K-medoids, also known as PAM or partitioning around medoids. K-means is the simplest and most commonly used method that follows an optimization approach by choosing the best solution from multiple runs. PAM, on the other hand, selects data points as centers or medoids and forms clusters using an arbitrary metrics of distances between these data points. K-means considers partitioning n observations into k clusters based on the minimized Euclidean distance, and PAM does the same but based on Manhattan distance. PAM is more robust than K-means when working with many outlier values since it is, unlike K-means, insensitive to outliers. Therefore, we employed PAM alongside with K-means to enable us to decide on which fits our data best. The following table shows the 3-cluster solution for both K-means and PAM:
Due to a large number of outliers in our data, PAM generated a more preferable and stable solution than K-means (Table 1) and accordingly was chosen for our study. PAM’s 3-cluster solution was validated using Silhouette method (Rousseeuw, 1987) reflecting a better placement of each observation in a cluster. Silhouette improves the results of cluster analysis, and its average width can be used to provide an assessment of cluster validity. The higher Silhouette coefficient, the better cluster validity. A positive number means observations are clustered properly; a negative number means observations could be placed in a wrong cluster, while zero means observations lie between two neighboring clusters. PAM’s 3-cluster solution had a positive Silhouette coefficient, which indicates our observations were clustered properly.

PAM’s 3-cluster solution is briefly explained below and visualized in figure 2:

**Cluster 1 – light users (n= 2087):** this group is the largest cluster and characterized by its low utilization of different MB services. As per figure 2 (the red area), the users in this cluster show a very low level of usage on all individual MB attributes than the other two clusters.

**Cluster 2 – moderate users (n= 1162):** this group is the middle cluster and characterized by moderate utilization of various MB services. As per figure 2 (the green area), the users in this cluster show a modest level of usage on all individual MB attributes.

**Cluster 3 – heavy users (n= 489):** this group is the smallest cluster and characterized by its high utilization of different MB services. As per figure 2 (the blue area), the users in this cluster show a higher level of usage on all individual MB attributes than the other two clusters.

<table>
<thead>
<tr>
<th></th>
<th>1st Cluster</th>
<th>2nd Cluster</th>
<th>3rd Cluster</th>
</tr>
</thead>
<tbody>
<tr>
<td>K-means</td>
<td>2965</td>
<td>712</td>
<td>61</td>
</tr>
<tr>
<td>PAM</td>
<td>2087</td>
<td>1162</td>
<td>489</td>
</tr>
</tbody>
</table>

![Figure 2. PAM’s 3-Cluster Solution](image)
Phase Two: Behavioral Study

The log data we use in the phase one does not contain any individual identifiers. However, there is a de-identified but unique consumer registrant number corresponding each customer. In phase one, we determined clusters based on this anonymized registrant number. The survey was sent by the bank to the same customers who were in the log file which has their registrant number and their email addresses. We did not have access to the email or other identifying information thereby satisfying the requirement of our IRB authorities. The surveys will be collected by the bank and data will be sent to us only with our de-identified registrant number for matching users from the log data file. This process allows us to uniquely match each customer between the log and survey data without compromising the individual privacy. We expect our sample size to reduce because all users may not respond. However, even if we expect a return rate of 20% we would have 748 users in our sample for the phase two of our study. We also expect a higher response rate from heavy users than moderate who will respond more than light users. This could balance the three clusters more evenly balanced for the upcoming structural model estimation.

We plan to study the matched data with a descriptive analysis of the demographic profiles by each user segment and analyze the factors of IS success model using structural equation modeling (SEM) technique to reveal the significant relationships with path coefficients of the tested model. Confirmatory factor analysis (CFA) will be employed to evaluate reliability through factor loadings, Cronbach's alpha, and composite reliability. Convergent validity will be checked using average variance extracted (AVE) while discriminant validity will be checked through comparing the square root of AVEs with other variables coefficients. Collinearity between variables will be assessed via variance inflation factor (VIF). Common method variance will be tested with Harman’s single-factor test (Zhou, 2012). We also plan to explore how the three homogenous segments differ on the influences of three quality factors on satisfaction and use and evaluate the goodness-of-fit measures. Finally, a multivariate analysis of variance (MANOVA) test will be used to explain the differences between group means of all constructs of our research model.

Conclusion and Possible Contributions

This research-in-progress presents an innovative data analytics approach to study MB system usage making theoretical and practical contributions. We combine objective with subjective measures to overcome the limitations of IS usage studies that only use subjective measures (Eckhardt et al. 2013) and use data analytics to segment our sample to overcome the bias from heterogeneous samples (Becker et al. 2013) both of which improve the validity of our study. A preliminary cluster analysis has grouped our users into three homogenous segments: light, moderate and heavy users. These segments were formed by using rich to lean MB usage measures from the system log files. Our next step is field study with these user segments using factors from the IS success model. We are using a well-established data analytics approach on a large user sample and IS theory on the system usage considering multiple factors and demographics influencing MB usage. While the analysis is work in progress, our preliminary segmentation results look promising. One limitation was our cluster sizes were uneven; we are exploring alternate clustering techniques like K/D tree for a better cluster distribution. Another limitation is that our findings cannot be generalized to all banks because our sample was from a mid-sized urban bank. A key challenge will be linking and integrating survey and log file data while protecting user privacy. Our results will provide the demographic profiles of three user groups, like their gender, age, education, and income, which when combined with MB success factors and usage data from log files provide a more accurate understanding of system success. The next phase of our analysis will hopefully reveal a higher level of explanatory power of the IS success model when considering the three segments separately rather than as one group.

Theoretical contributions from this study will be the application of IS success model for MB usage and segmenting users into groups based on their MB system usage behavior with a data analytics approach. To the best of our knowledge, the system usage variable has never been used to categorize users to study IS system success. Our data analytics approach could be used by other researchers for a better understanding of IS usage studies to reduce bias in parameter estimation and inferential errors through segmentation. Also, it will enable researchers to better understand the factors influencing usage (Becker et al. 2013), improve MB usage by considering individual needs (Schacht et al. 2015) and increase the overall system use through personalization (Bartis and Mitev, 2008). This latter theoretical contribution also has practical contributions. Segmentation will help banks make better decisions on customizing system features when
combined with the rich demographic data about each user group, like their gender, age, education and income information from the survey Banks can use our study outcomes for supporting more specific system features that are tailored for smaller user groups, increasing customer loyalty and bank revenue. This we believe will increase MB system usage and increase the chances for system success.

Acknowledgement

We would like to thank Stephen Irish and Brian Collins of the Enterprise Bank in Lowell for collaboration on this project. They provided us with the MB system log files and conducted survey with the bank customers for this project.

References


