12-13-2018

An Exploratory Study of Information Systems Usage Profiles from a Longitudinal Perspective

Khawaja A. Saeed
Wichita State University, khawaja.saeed@wichita.edu

David Xu
City University of Hong Kong, davidxu@cityu.edu.hk

An Exploratory Study of Information Systems Usage Profiles from a Longitudinal Perspective

Research-in-Progress

Khawaja A Saeed
Wichita State University
khawaja.saeed@wichita.edu

David Xu
City University of Hong Kong
davidxu@cityu.edu.hk

Abstract

Past research on Information Systems (IS) usage has mainly focused on the presence of usage and time of usage from a cross-sectional perspective, which provides limited understanding about what usage profiles exist and how these profiles change over time. We address this gap by leveraging a data set that captures IS usage profiles from IS usage log data over 18 months. Through cluster analysis based on two dimensions (number of IS features used and the depth of such usage), we found that the longitudinal profiles represent five distinct trajectories: minimalists, centrists, maximizers, decliners, expanders. Expanders start with a low usage profile and expand the usage over time, while all other four groups started the usage at a medium or high usage level and then made different adjustments over time. Moving forward it will be interesting to evaluate cognitive and behavioral differences across these profiles.

Keywords: Information systems usage profiles, longitudinal cluster analysis, mobile banking

An Exploratory Study of Information Systems Usage Profiles from a Longitudinal Perspective

Research-in-Progress

Introduction

Information systems (IS) usage has been examined as a situated or temporal variable that can assume a varying level of richness in conceptualization and measurement. It is logical to assume that higher usage which has been defined in different ways is a normative objective tied to better performance outcomes (Limayem and Hirt 2003). Despite the extensive research on IS usage, our understanding of IS usage at the post adoption phase is rather limited. Especially, few prior studies have evaluated IS usage from a longitudinal perspective based on objective data. Users may manifest certain profiles based on their interaction with an IS over time. The longitudinal profiles represent distinct trajectories that provide insights on the temporal evolution of IS usage behavior. It is unclear whether such trajectories exist. If so, how many and what is their structure. Limited work has evaluated these questions through direct measurement of IS usage in a natural setting. This study addresses these questions.

Jasperson et al. (2005) highlight the need to identify patterns of IS use and how these patterns evolve over time. The work by Burton-Jones and Straub (2006) structures the system usage phenomenon on a lean vs rich continuum. They propose that usage needs to be examined from a system, user, and task perspectives with richer measures capturing more elements (system, user, and task). In practice, IS usage is associated with user types. For example, "super-user" is a label that is commonly used. The richness of IS usage is reflected in the behavior of the super-user. Thus, an alternative perspective is to view IS usage from a profile perspective which is reflected in how users interact with the IS. For example, prior research on IS usage highlights profiles such as benefits maximization, benefits satificing, disturbance handling, and self-preservation (Beaudry and Pinsonneault, 2005). A key characteristic of these usage types and profiles is that they capture user's interaction with the IS at a moment in time.

A longitudinal profile that captures the IS – User interaction over time provides an alternative perspective. For example, if we observe IS usage behavior of multiple users repeatedly over time, a sub group of users may show an enhanced usage profile where in after initial adoption this group shows an increasing trajectory by expanding the level of IS usage. Another sub group may show a decreasing trajectory by initially starting at a high level of usage and minimizing IS usage over time. We refer to these profiles as evolutionary trajectories. In order to get insights on these trajectories, user behavior needs to be observed from the time the system is introduced to the potential users to over a substantial period of usage.

Two contrasting perspectives exist on IS usage profiles based on evolutionary trajectories (Ortiz de Guinea and Webster, 2013). Prior studies provide strong evidence to support the concept of habitual usage (Limayem and Hirt 2003). Once a user develops a stable IS usage pattern, inertia kicks in and creates strong resistance to change in behavior. On the contrary, the adjustment perspective argues that individuals are active users and consistently evaluate their expectations from the IS and reflect on their usage experience (Kim and Malhotra, 2005). The assessment can trigger adjustments in future usage behavior (Kim, 2009). Adjustments may also require formal triggers / events initiated by another entity or external disruption emanating from the task environment (Sun, 2012). Investigation of evolutionary trajectories with respect to IS usage can help in providing insights about these contrasting perspectives regarding which one is more likely to hold.

The study examines two interesting but unexplored questions. First, what evolutionary trajectories develop after the introduction of an information system in a voluntary usage setting? Second, are these trajectories persistent or do they change over time and if so how? We evaluate these questions using a data set that captures objective IS usage in a natural setting over an extensive period of time.

Literature Review

Ortiz de Guinea and Webster (2013, p.1166) defines IS use patterns as "the configuration of an individual's emotions (what the user feels), cognitions (what the user thinks), and behaviors (what the user does) when interacting with IT to accomplish a task". Our study focuses on configuration of user behaviors by capturing the user's interaction with an IS at the feature level for task accomplishment. Identifying behaviors that can provide insights on the variation in user system interaction can help in profile development. For example, Beaudry and Pinsonneault (2005) through interview data segmented users into different post adoption strategies based on the extent to which they feel in control and view the IS as an opportunity / threat. They categorized the users into benefits maximization, benefits satificing, disturbance handling, and self-preservation and associate these categories with varying IS usage patterns. Prior research has predominately focused on IS usage dimensions (see the literature review in Table 1). The discussion has centered on comparing the dimensions in their efficacy to explain usage continuance or performance outcomes. We argue that IS usage dimensions when considered jointly can assist in identification of user profiles.

Table 1: Review of IS Usage Behavior

Cite	IS Usage Perspectives
Lassila and Brancheau, 1999	Volume of IT usage (Number of Hours), Diversity of IT usage (# of features used), Reliance on IT (dependence on IT to get the job done)
Devaraj and Kohli, 2003	Total number of DSS reports accessed, Processing time (CPU), Number of records accessed
Ahuja and Thatcher, 2005	Trying to Innovate with IT
Burton-Jones and Straub, 2006	Time of usage, Cognitive absorption, Deep structure usage
Sun, 2012	Try new features, Feature substituting, Feature combining, Feature repurposing
Bagayogo et al., 2014	Locus of Innovation, Substantive use, Adaptive use

Once a user develops a profile, does the profile remain situated or does it change over time? Prior research offers two perspectives related to this question. Automaticity perspective envisions the development of a stable use pattern that persists over time. While the adjusting perspective argues for adaptation in use of IS over time but points towards triggers and events that disrupt existing usage or create an incentive for change in usage behavior. The adaptive perspective emphasizes that users can choose to increase usage, decrease usage, or completely abandon usage (Sun, 2012). Maintaining status quo is also recognized as a choice and is reflective of the recognition that automaticity is a possibility.

The review shows that using longitudinal objective IS usage data to uncover IS usage profiles can provide a significant contribution. Investigating whether these profiles evolve or remain persistent over time can assist in assessing the validity of automaticity vs. adjustment paradigms. Our study relies on an inductive approach to examines these issues because prior research offers limited guidance on the evolutionary trajectories of IS usage.

Data Collection and Research Methods

The data on actual mobile banking usage from the launch (April 2008) of the mobile banking service to October 2009, a total of 18 months were obtained from a financial institution in the mid-west region of the United States. The dataset consisted of transactional data of mobile banking system usage (it did not include interaction with the online banking system through a computer). It was ascertained that users in the data set remained customers during the 18 month period. Users that opened a new account within the 18 months period were retained in the data because if the account was opened 3 months prior to October

2009, this provided the minimum number of data points needed to conduct the analysis. Due to data privacy regulations and requirements, we used scripts to anonymize the dataset but kept track of the users over the 18 months by using alternative identifiers. The dataset captured the date and time when a user logged in, description of features used in that session, time spent on using each feature, session ID to capture an interaction episode, mobile device used, and when the user logged out. The Data cleaning and aggregation required a substantial effort. The month was selected as the unit of analysis and data were aggregated to capture the usage parameters on a monthly basis. The core reason for selecting the month to aggregate the data was that it aligns well with how individuals managing their financials. Individuals manage their income and expenses on a monthly basis which is directly related to their interaction with the banking system. Short terms financial objectives and obligations are normally structured on a monthly basis as well.

IS Usage Parameters

As discussed earlier, our objective was to capture the user's interaction with the system in a comprehensive manner. We selected two measures that capture different perspectives on IS usage. Time of IS usage has been used in multiple prior studies (Burton-Jones & Straub, 2006). However, this measure does have a limitation that it does not effectively capture the purpose of usage. Another limitation which raises concerns about the validity of the measure is that unless explicitly observed it is not possible to ascertain if the user was actively interacting with the system. Although, this measure was available, due to these limitations we decided against using it. We selected the number of features used by the user to capture IS usage. From the perspective of the richness of IS usage, this parameter offers a better approach because it captures the user, task, and the system. We term this parameter feature usage. Although feature usage offers a better approach it does not distinguish between features that may be considered basic or advanced. Thus, the second parameter we captured was the depth of IS feature usage. This parameter measured the highest depth of IS feature (ranging from basic to advanced) usage achieved by the user.

Feature Usage = Number of features used in a month

Feature Depth = Highest depth of feature use (basic to advanced) in a month.

The feature descriptions captured whether the user accessed account information, reviewed account history, used bill pay, and / or transferred funds. Feature usage measured the number of features a user used in a month. We segmented usage features into three levels of usage: (1) account information, (2) account history, and (3) bill pay / funds transfer, with (1) representing basic feature usage and (3) representing the most advanced feature usage. Feature depth was computed by the highest level of depth achieved in a month. For example, in a session, if the user accessed account information and also used bill pay, two features were used and the feature depth was coded as 3. Therefore, feature usage was the sum of features and feature depth only captured the highest level of feature used in a session. If the user had 10 sessions in a month, the number of features were aggregated over that time period and the level of feature depth was coded based on the highest level of feature used in those sessions. Table 2 provides aggregate mean and standard deviation for the two variables over an 18 months period.

Table 2: Descriptive Statistics

Variables	Minimum	Maximum	Mean	S.D.
Feature Usage	1.00	105.31	7.50	10.53
Depth of Feature Usage	1.00	3.00	1.87	0.69

Evolution of IS Usage Profiles

Ortiz de Guinea and Webster (2013) show a cyclical pattern in which a user may oscillate between automaticity and adjusting behaviors driven by the type of event. However, automaticity is shown to be the dominant usage pattern. They used two experiments. The first experiment was conducted over two weeks and the second experiment was based on a single interaction episode that lasted 1.5 hours. No prior study to our knowledge has examined IS usage profile development and evolution in a setting in which a new system was introduced to the users. It will be interesting to get insights on the time it takes for the users to adopt a specific usage pattern. Do the users quickly settle into a stable usage pattern and if so how long does it persist? Do the users oscillate between automatic and adjusting behaviors?

In order to analyze evolutionary trajectories and explore the answer to these questions, we used Kml3d package in R that uses k-means clustering in the context of longitudinal data. Kml3d is designed to cluster several variables – trajectories jointly. For example, multiple attributes related to a phenomenon can be jointly used to identify clusters in a longitudinal dataset. Kml3d also offers various quality indices to support decision making regarding the appropriate number of clusters and supports methods to handle missing data. The details about the mathematical approach used in Kml3d and its implementation can be found in Genolini et al. (2015) and at https://cran.r-project.org/web/packages/kml3d/index.html.

Kml3d supports the capability to develop and view 3D graphs of the results and provides criteria such as Calinski & Harabasz, Calinski, Harabasz, Kryszczuk variant, Calinski, Harabasz, Genolini variant, Ray & Trui, Davies & Bouldin, BIC, and AIC as decision aids for selecting the appropriate number of clusters (Genolini et al., 2015). The Calinski variations use between cluster covariance and within cluster covariance to measure the compactness of clusters with some variations. Ray & Tui and Davies and Bouldin use the distance between two cluster points and the distance between cluster centers for computation of the criterion. These criteria are non-parametric measures. Bayesian Information Criterion (BIC) and Akaike Information Criterion (AIC) are also available as parametric criteria. Kml3d computation for these criteria is such that high values show better partitioning (Genolini, et al., 2015). Kml3d automatically normalizes the data for analysis so the results are not impacted by scale effect. Furthermore, after analyzing the user can output cluster association for further assessment. The R-code we used for Kml3d is provided in Appendix A.

Results

We configured kml3d to provide us with 3 to 5 cluster solutions and run the k-means for k= 3, 4, 5, 50 times each using all methods. Genolini and Falissard (2011) based on simulated data report that all methods option combines the different starting conditions and provides the most efficient method. As reported by Genolini and Falissard (2011), Hartigan and Wong (1979) version of k-means is used and the default distance is Euclidean with Gower adjustment. The data needs to be transformed into "ClusterLongData3D" object using the function cld3d from data.frame so that it is appropriately configured for analysis. Data was available from when the system was introduced to the users to 18 months after initial release. Not all users adopted the system at initial release, so data configuration posed challenges. In order to develop trajectories, we need to capture the data from when the user initially adopted the system and the user's subsequent usage. The approach that we used was to code the data as NA (not available) if the user had not adopted the system in that month yet. After the adoption month if the user did not use the system, we coded the usage as zero. In the specification of the parameters for Kml3d, we specified that a minimum of 3 data points should be available for the user to be included in the analysis. That is, the user should have interacted with the system for at least three months to be included in the analysis.

We present the results for all the cluster solutions and offer discussion on which solution may be appropriate. The three clusters solution segments the users into A, B, and C clusters (Figure 1 & 2). Users in group A start their interaction with mobile banking at the medium levels of feature usage and depth but over time shift to lower levels of usage across both dimensions. In the first 5 months, there is a consistent decrease in feature usage and feature depth, a slight increase from month 5 to month 6 and afterward there is a gradually decreasing trend on both usage types. We call users in group A minimalists. Group B initiated the interaction with mobile banking at a high level, but adjusted their usage slightly to the downside. The first 6 months show an oscillating pattern. After that the users in group B, settled into a pattern that shows average levels of feature usage and depth for the next 4 months. After month 10, these users depict a gradual decrease in both feature usage and feature depth and then settle into a stable pattern in the last two month. We call users in group B centrists. Users in Group C started with a pattern that was similar to group B. They started using the system in month 2 which is evident in the lack of a trajectory in the first month. There is a drop in usage across both dimensions until month 5. Afterwards, there is a consistent pattern of increase in usage across both features and depth except for the time frame between month 14 and 15. We call the users in group C maximizers. Figure 2 provides a 3D view of the clusters jointly across the two usage dimensions. In terms of classification, 70% of the users were categorized as minimalists, 26% were categorized as centrists, and 4% were categorized as maximizers.

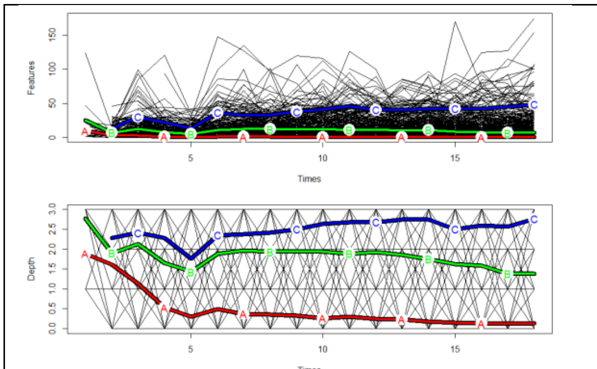

Figure 1: Three Clusters Solution across IS Usage Parameters

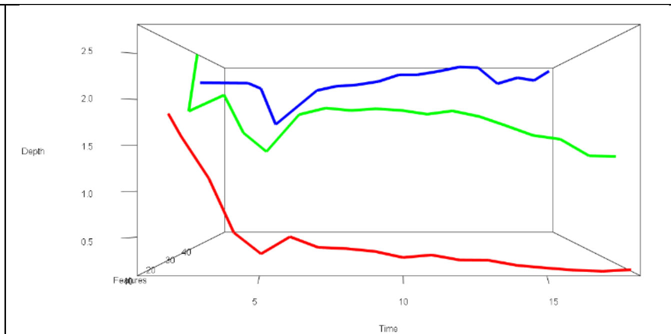

Figure 2: 3D graph for Three Clusters Solution

The four clusters solution has the same three clusters identified in the three clusters solution. The additional cluster looks like a variation of minimalists and centrists in the three clusters solution (Figure 3 and Figure 4). The new cluster shows a usage pattern that is similar to centrists in the initial months of usage. However, after month 6 there is a gradual decrease in both feature usage and depth of usage. In the last three month, the new group in the four clusters solution depicts a usage pattern that becomes very similar to minimalists. We call the new group decliners. The classification shows that 56% of the users were categorized as minimalists, 25% of the users were categorized as decliners (new cluster), 15% of the users were categorized as centrists (this was cluster B in the three clusters solution), and 4% were categorized as maximizers (cluster C in the three clusters solution). The new cluster draws 14% of the members from minimalists (group A in the three clusters solution) and 11% of the members from centrists (group B in the three clusters solution).

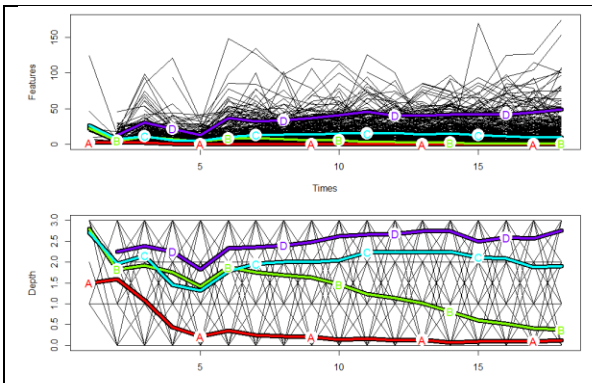

Figure 3: Four Clusters Solution across IS Usage Parameters

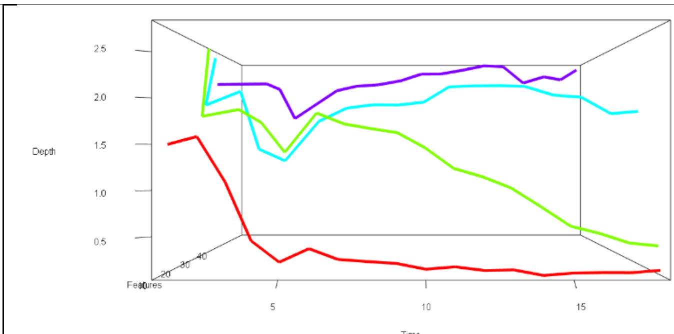

Figure 4: 3D graph for Four Clusters Solution

The five clusters solution has the same four clusters that were found in the four 4 clusters solution. The new cluster shows a pattern in which the user initiates the interaction with mobile banking at a relatively low level of feature usage and depth (Figure 5 and Figure 6). After month 5 there is a significant increase on both IS usage dimensions. After month 14, the new group identified in the five clusters solution has the highest level of usage in terms of features and depth. We call this group expanders. The classification shows that 56% were categorized as minimalists, 25% as decliners, 13% as centrists, 4% as maximizers, and 2% as expanders (new group / group E). Expanders do provide a unique pattern. However, it only represents 2% of the users. These users represent a combination of users that were categorized as decliners and centrists in the four clusters solution.

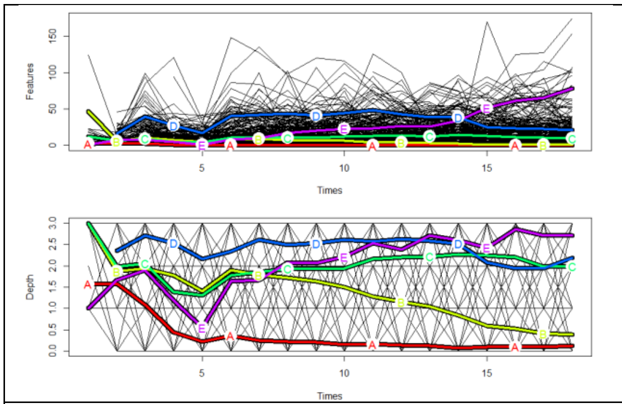

Figure 5: Five Clusters Solution across IS Usage Parameters

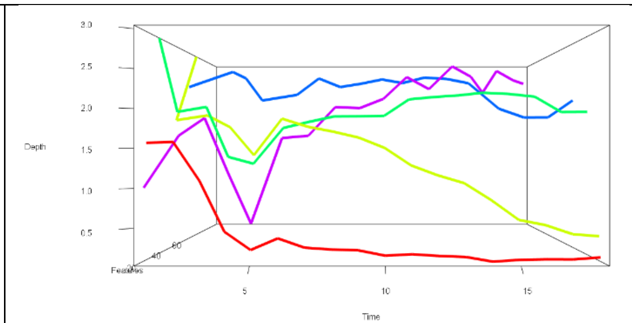

Figure 6: 3D graph for Five Clusters Solution

Table 3 provides information on the actual number of users in each cluster across the three different cluster solutions. Table 4 provides the average for feature usage and depth of usage across the 18 months. It should be noted that the important aspect to consider in separation of the clusters is the trajectories.

Table 3: Cluster Membership

Clusters	Minimalists (Group A)	Decliners (Group B)	Centrists (Group C)	Maximizers (Group D)	Expanders (Group E)
Solution 1	916		332	55	
Solution 2	722	335	194	52	
Solution 3	722	314	195	49	23

Table 4: Means across Clusters (18 months average)

Clusters / Usage Dimensions		Minimalists (Group A)	Decliners (Group B)	Centrists (Group C)	Maximizers (Group D)	Expanders (Group E)
Solution 1	**Feature Usage**	3.76		12.22	43.07	
	Depth of Usage	1.69		2.26	2.68	
Solution 2	**Feature Usage**	3.35	6.74	14.96	44.17	
	Depth of Usage	1.61	2.04	2.38	2.68	
Solution 3	**Feature Usage**	3.40	6.93	12.19	38.60	42.38
	Depth of Usage	1.62	2.02	2.30	2.64	2.55

Table 5 provides results for the different selection criteria. We evaluated both nonparametric criteria and parameter criteria. There was variation across the different criteria in terms of the best cluster solution.

As suggested by Genolini et al., (2015), it is important to balance prior knowledge about the existence of user groups, cluster quality criteria, and solution practicality in choosing the cluster solution. The results based on quality criteria support either a three or a five clusters solution. The parametric criteria (BIC and AIC) and a variation of the Calinski and Harabatz support the existence of five clusters. The three clusters solution is probably cleaner. However, the five clusters solution does show distinct trajectories when observed over the entire usage time frame.

Table 5: Selection Criteria

Solutions / Criteria	Calinski. Harabatz	Calinski. Harabatz2	Ray.Turi	Davies. Bouldin	BIC	AIC
Solution 1	**450**	0.70	-0.04	**-1.50**	-79000	-79000
Solution 2	350	0.80	-0.04	-1.65	-77000	-76500
Solution 3	275	**0.85**	-0.04	-1.65	**-76000**	**-75000**

Discussion

The analysis highlights three interesting insights. First, the notion that users start with a low usage profile and expand the usage over time is not supported by the results. Except for one group (i.e., expanders) in the five clusters solution which was also the smallest cluster in terms of size, all other groups actually started the usage at a medium or high usage level and then made adjustments overtime. It is also evident that the first five months show higher variation in usage indicating experimentation. The users moved to a more distinct and stable usage pattern after that time frame.

Second, the composition of the clusters shows that the largest number of users are in a cluster that depicts the lowest level of feature usage and usage depth. Furthermore, the cluster with the second largest membership although initially shows a pattern of usage that is high on both usage dimensions, overtime the usage drops significantly. Approximately 80% of the users fall in either minimalists or decliners categories and the remaining 20% were categorized as centrists, maximizers, and expanders. These groups differed in the initial month after the introduction of the system but over time developed a trajectory that shows relatively high level of usage across both usage dimensions (features and depth).

Third, the results need to be examined in the light of what earlier studies have assessed related to this topic. Prior research emphasizes the importance of understanding usage at the feature level and this study builds on that concept. Earlier studies have addressed adaptive usage and automaticity of use as the dominant pattern but also highlighted the possibility of oscillation between automatic and adjusting behaviors. Our study offers the first preview of evolution of IS usage behavior over an extended period of time. It does mirror to some extent situated usage patterns such as low and enhanced usage but confirm the persistence of such patterns of use over time. Other patterns are unique and can be understood only when examined from a longitudinal perspective. For example, decliners and expanders capture adaptive behavior over time.

We intend to extend this work by evaluating the relationship between the IS usage evolutionary trajectories and continued IS usage. It will be interesting to assess if some profiles depict a better chance of continuing to use the system. Assessment of trajectories also offer a different perspective on understanding the concept of adaptive usage.

References

Ahuja, M. K. & Thatcher, J. B. 2005. Moving beyond intention toward theory of trying: Effects of work environment and gender on post-adoption information technology usage, *MIS Quarterly*, (29:3), pp. 427-459.

Bagayogo, F. F., Lepointe, L., & Bassellier, G. 2014. Enhance IT use: A new perspective on post-adoption, *Journal of the Association of Information Systems*, (15:7), pp. 361-387.

Beaudry, A., & Pinsonneault, A. 2005. Understanding user responses to information technology: A coping model of user adaptation. *MIS Quarterly,* (29:3), 493-524.

Burton-Jones, A., & Straub, D. W. 2006. Reconceptualizing system usage: An approach and empirical test. *Information Systems Research,* (17:3) 228-246.

Devaraj, S. & Kohli, R. 2003. Performance impacts of information technology: Is usage the missing link?, *Management Science*, (49:3), pp. 273-289.

Genolini, C. 2015. kml3d: K-Means for Joint Longitudinal Data. R package version 2.3, URL http://CRAN.R-project.org/package=kml3d.

Genolini, C. & Falissard, B. 2011. Kml: A package to cluster longitudinal data, *Computer Methods and Programs in Biomedicine*, 2011.

Genolini, C., Alacoque, X., Sentenac, M., & Arnaud, C. 2015. Kml and Kml3d: R packages to cluster longitudinal data, *Journal of Statistical Software*, (65:4).

Hartigan, J., & Wong, M. 1979. A K-means clustering algorithm, *JR Stat. Soc. Ser C-Appl Stat*, 28, pp. 100-108.

Jasperson, J., Carter, P. E., & Zmud, R. W. 2005. A comprehensive conceptualization of post-adoptive behaviors associated with information technology enabled work systems, *MIS Quarterly,* (29:3), pp. 525-557.

Sun H. 2012. Understanding user revisions when using information system features: Adaptive system use and triggers, *MIS Quarterly*, (36:2), pp. 453-478.

Kim, S. S. 2009. The integrative framework of technology use: an extension and test, *MIS Quarterly*, (33:3), pp. 513-537.

Kim, S. S., & Malhotra, N. K. 2005. A longitudinal model of continued IS use: An integrative view of four mechanisms underlying postadoption phenomena, *Management science*, (51:5), pp. 741-755.

Lassila, K. S., & Brancheau, J. C. 1999. Adoption and utilization of commercial software packages: Exploring utilization, equilibria, transitions, triggers, and tracks, *Journal of Management Information Systems*, (16:2), pp. 63-90.

Limayem M, & Hirt, S. G. 2003. Force of habit and information systems usage: Theory and initial validation, *Journal of the Association for Information Systems*, (4:1), pp. 65-95.

Ortiz de Guinea, A., & Markus, M. 2013. An Investigation of Information Systems Use Patterns: Technological Events as Triggers, the Effect of Time, and Consequences for Performance, *MIS Quarterly*, (37:4), pp. 1165-1188.

Appendix A: R Code for Kml3d

library(kml3d) –load kml3d package
Ranalysis <- read.csv(file="filename.csv", head=TRUE, sep=",") – load the data into a dataframe
anaCLD <- cld3d(Ranalysis, timeInData = list(Features = 2:19, Depth=20:37), maxNA =15) – load the data into a cld3d data object. The data file should be set up so that the column headings reflect longitudinal data. For example, there are eighteen columns for features and they were labeled as Feature1, Feature2, Feature3….
kml3d(anaCLD, 3:5, nbRedrawing = 50, parAlgo = parKml3d(centerMethod = meanNA, startingCond = "all", scale=TRUE)) – run the kml3d cluster analysis for 3 – 5 cluster solutions
try(choice(anaCLD)) – display the results for the cluster solutions
plotMeans3d(anaCLD, 5) – 3D chart for a specific cluster solution
myclusters <- getClusters(anaCLD, 5) – store cluster association for users into a data object
write.csv(myclusters, file = "myclusters.csv") – write the cluster association to a csv file