Modeling Long-Term Search Engine Usage

Ryen W. White, Ashish Kapoor, and Susan T. Dumais

Microsoft Research,
One Microsoft Way, Redmond WA 98052, USA
{ryenw,akapoor,sdumais}@microsoft.com

Abstract. Search engines are key components in the online world and the choice of search engine is an important determinant of the user experience. In this work we seek to model user behaviors and determine key variables that affect search engine usage. In particular, we study the engine usage behavior of more than ten thousand users over a period of six months and use machine learning techniques to identify key trends in the usage of search engines and their relationship with user satisfaction. We also explore methods to determine indicators that are predictive of user trends and show that accurate predictive user models of search engine usage can be developed. Our findings have implications for users as well as search engine designers and marketers seeking to better understand and retain their users.

Keywords: Search Engine, Predictive Model.

1 Introduction

Search engines such as Google, Yahoo!, and Bing facilitate rapid access to the vast amount of information on the World Wide Web. A user's decision regarding which engine they should use most frequently (their *primary* engine) can be based on factors including reputation, familiarity, effectiveness, interface usability, and satisfaction [14], and can significantly impact their overall level of search success [18, 20]. Similar factors can influence a user's decision to switch from one search engine to another, either for a particular query if they are dissatisfied with search results or seek broader topic coverage, or for specific types of tasks if another engine specializes in such tasks, or more permanently as a result of unsatisfactory experiences or relevance changes, for example [18]. The barrier to switching engines is low and multiple engine usage is common. Previous research suggests that 70% of searchers use more than one engine [20].

Research on engine switching has focused on characterizing short-term switching behavior such as predicting when users are going to switch engines within a search session [7, 10, 18], and promoting the use of multiple engines for the current query if another engine has better results [20]. Given the economic significance of multiple engine usage to search providers, and its prevalence among search engine users, it is important to understand and model engine usage behavior more generally. Narrowly focusing on the use of multiple search engines for a single query or within a single search session provides both limited insight into user preferences and limited data

P. De Bra, A. Kobsa, and D. Chin (Eds.): UMAP 2010, LNCS 6075, pp. 28–39, 2010.

from which to model multiple engine usage. There has been some research on modeling engine usage patterns over time, such as studies of switching to develop metrics for competitive analysis of engines in terms of estimated user preference and user engagement [8] or building conceptual and economic models of search engine choice [13, 17]. Rather than characterizing and predicting long-term engine usage, this previous work has focused on metric development or has specifically modeled search engine loyalty.

In this paper we model user behaviors and determine key variables that affect search engine usage. In particular, we study engine usage of over ten thousand consenting users over a period of six months using log data gathered from a widely-distributed browser toolbar. We perform analysis of the data using machine learning techniques and show that there are identifiable trends in the usage of search engines. We also explore methods to determine indicators of user trends and show that accurate predictive user models of search engine usage can be developed. Knowledge of the key trends in engine usage and performant predictive models of this usage are invaluable to the designers and marketers of search engines as they attempt to understand and support their users and increase market share.

The rest of this paper is structured as follows. Section 2 presents related work on search engine switching, engine usage, and consumer loyalty/satisfaction. Section 3 presents our research and experiments on modeling and predicting engine usage. Section 4 discusses our approach and its implications. We conclude in Sect. 5.

2 Related Work

The most significant related work lies in the areas of search engine switching, consumer choice regarding search engine usage, and studies of consumer loyalty with a product or brand and their associated levels of satisfaction. In this section we describe work in each of these areas and relate it to the research described in this paper.

Some research has examined engine switching behavior within a search session. Heath and White [7] and Laxman et al. [10] developed models for predicting switching behavior within search sessions using sequences of user actions. White and Dumais [18] used log analysis and a survey to characterize search engine switching behavior, and used features of the active query, the current search session, and the user to predict engine switching. White et al. [20] developed methods for predicting which search engine would produce the best results for a query. One way in which such a method could be used is to promote the use of multiple search engines on a query-by-query basis, using the predictions of the quality of results from multiple engines. Studying switching within a session is useful for understanding and predicting isolated switching events. However, to develop a better understanding of users' engine preferences and model multi-engine usage effectively we must look beyond isolated sessions. The research presented in this paper leverages engine usage statistics aggregated weekly over a six-month period to build models of search engine usage across many thousands of search engine users.

Mukhopadhyay et al. [13] and Telang et al. [17] used economic models of choice to understand whether people developed brand loyalty to a particular search engine, and how search engine performance (as measured by within-session switching) affected user choice. They found that dissatisfaction with search engine results had both short-term and long-term effects on search engine choice. The data set is small by modern log analysis standards (6,321 search engine switches from 102 users), somewhat dated (data from June 1998 – July 1999 including six search engines but not Google), and only summary level regression results were reported. Juan and Chang [8] described some more recent research in which they summarize user share, user engagement and user preferences using click data from an Internet service provider. They identify three user classes (loyalists to each of the two search engines studied and switchers), and examine the consistency of engine usage patterns over time. We build on this work to identify key trends in search engine usage, reason about why certain behaviors were observed, and develop predictive models of search engine usage based on observed usage patterns and user satisfaction estimates.

There is a large body of work in the marketing community regarding product or brand switching and the relationship between satisfaction and loyalty. Research in these areas is typically concerned with identifying factors that influence customer defections [5] or developing models of satisfaction or loyalty that make it easier for businesses to understand customer rationale and take corrective action if needed [9, 11]. In this paper we model the usage of multiple search engines over time and base part of our model on estimates of searcher satisfaction gleaned from log data. Although satisfaction is the predominant metric used by companies to detect and manage defections to competitors [1,3], more recent research has found that knowledge of competitors and attitudinal and demographic factors, among other influences, can also play an important role [14].

To summarize, the research presented in this paper differs from earlier work in that we study patterns of usage for multiple search engines over a long period of time, identify key trends, and develop predictive models for these trends.

3 Modeling Search Engine Usage

Understanding and predicting retention and switching behavior is important for search engine designers and marketers interested in satisfying users and growing market share. To model long-term patterns of search engine use, we used data from a widely-distributed browser toolbar (described in more detail below). We examine the patterns of search engine usage for tens of thousands of users over a six-month period of time. We first identify key trends in people's usage patterns (e.g., sticking with the same engine over time, switching between engines, etc.). We then summarize features that distinguish among the different usage trends. Finally, we develop models to predict which trend a particular individual will follow over time.

Data Collection. We used six months of interaction logs from September 2008 through February 2009 inclusive, obtained from hundreds of thousands of consenting users of a widely-distributed browser toolbar. These log entries include a unique identifier for the user, a timestamp for each Web page visited, and the URL of the Web page visited. Intranet and secure (https) URL visits were excluded at the source. To remove variability caused by geographic and linguistic variation in search behavior, we only include log entries generated in the English speaking United States locale.

From these logs, we extracted search queries issued to three popular Web search engines (which we call A, B and C) for each of the 26 weeks. We selected users who issued at least 10 queries per week in all weeks of the study. Applying this threshold gave us a sufficient number of queries each week to reliably study engine usage. These users formed the pool from which study subjects were randomly selected and for whom engine usage models were constructed.

Features. We extracted several different features to describe user interaction with the search engines. These features are shown in Table 1. We summarize the proportion of queries issued to each search engine. We also use the number of queries and the average length of the queries sent to each engine. Finally we use measures of a user's satisfaction with the search engine. Satisfaction estimates for each of the engines are determined using the fraction of queries issued to that engine that have post-query dwell times equaling or exceeding 30 seconds. This estimate of user satisfaction is based on results from Fox et al. [6], in which they learned models of which implicit user interactions (such as page dwell time) are predictive of explicit judgments of user satisfaction. We further broke down satisfaction according to whether queries were navigational or not because poor performance on this usually straightforward class of queries may be especially likely to result in engine switches. Navigational queries were defined as queries for which the same search result was clicked at least 95% of the time. For each user we have three values for each of these features (one corresponding to each search engine) for each of the 26 weeks of our study.

Table 1. Features of user interaction with each search engine (per week)

Feature	Description
fractionEngine	Fraction of queries issued to search engine
queryCountEngine	Number of queries issued to search engine
avgEngineQueryLength	Average length (in words) of queries to search engine
fractionEngineSAT	Fraction of search engine queries that are satisfied
fractionNavEngine	Fraction search engine queries defined as navigational
fractionNavEngineSAT	Fraction of queries in fractionNavEngine satisfied

3.1 Key Trends in Long-Term Search Engine Usage

We first analyze the data using dimensionality reduction techniques to develop insights about key trends. In particular, we use non-negative matrix factorization (NNMF) [10] of the data to summarize key behavioral patterns using a small number of basis vectors. NNMF is a method to obtain a representation of data using non-negativity constraints. This leads to part-based representation such that the original data can be represented as an additive combination of a small number of basis vectors.

Formally, we construct $\mathbf{A}$, an $M \times N$ matrix, where the columns represent users (N= 10,000 users) and the rows are the vector representation of 26 weeks of search engine usage statistics. The observations we used for our initial analysis are the proportion of queries that each user issued to each of the three search engines, thus $M = 26 \times 3$.

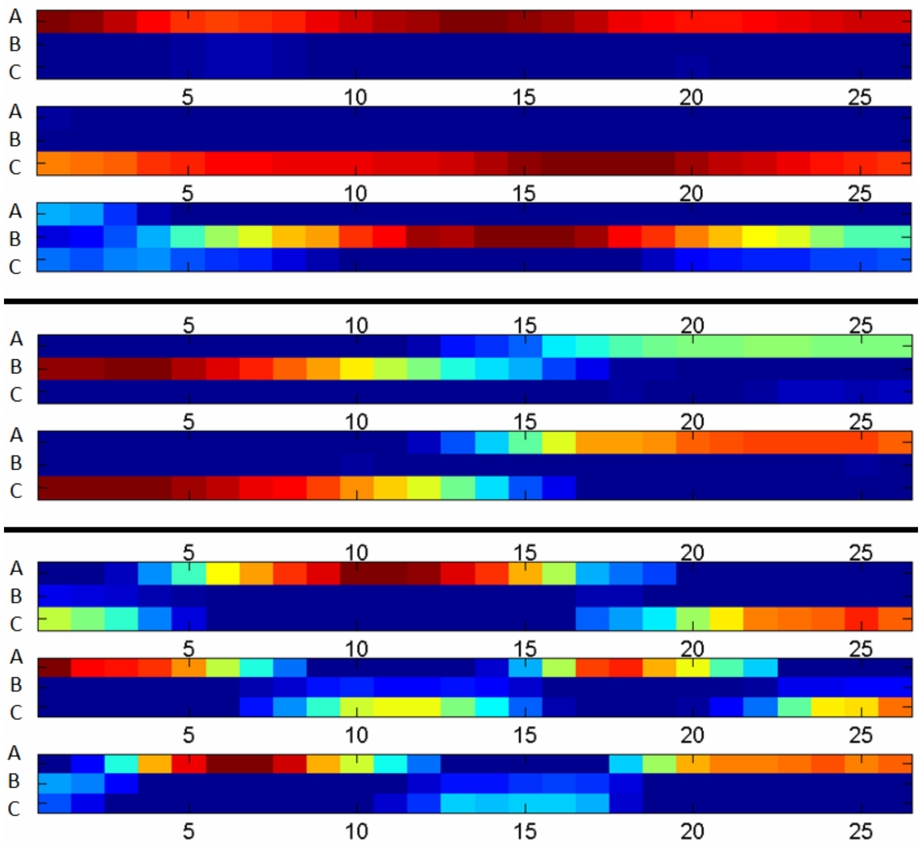

Fig. 1. Basis vectors learned from non-negative matrix factorization. We see three main patterns: sticking to a particular engine (first, second and third rows), switches that persist (fourth and fifth rows) and oscillations between different choices (sixth, seventh and eighth rows).

This matrix is decomposed into non-negative vectors matrix **W** and **H** of dimensions $M \times R$ and $R \times N$ respectively. Each column of **W** represents a basis vector while each row of the matrix **H** encodes the weight of the corresponding basis vector for each of the N users. The non-negative factorization means that we can interpret each individual observation vector as an additive combination of basis vectors (or parts). Consequently, each of the basis vectors represents the key trends that make up the behaviors that we observe in all the users.

Figure 1 shows the results of this analysis where we depict the top eight recovered basis vectors. Each image in the figure depicts one vector comprising the 26 weeks (x-axis) usage of three search engines (y-axis, labeled A, B, C). The red end of the spectrum represents high usage of an engine and the violet end represents low usage. A light blue coloring represents roughly equal usage of all three engines. In the first row, for example, we see a group of users who consistently use search engine A throughout the six-month period we studied. The fifth row shows users who initially use search engine C almost all the time, gradually decrease their usage of C (color transition from red to light blue), and slowly increase their usage of engine A. The seventh row show a pattern of behavior in which search engine A is used initially, followed by a gradual move to engine C (although it is not used all of the time, as indicated by the green rather than red color), then back to A, and finally back to C.

In general, this NNMF analysis identifies three key behavioral patterns:

1) *No Switch*: Sticking to one search engine over time (first, second, third rows),
2) *Switch*: Making a switch to another search engine and then persisting with the new engine (fourth and fifth rows), and
3) *Oscillation*: Oscillating between different engines (sixth, seventh, eighth rows).

While a majority of users indeed stick to one particular search engine, from the user modeling perspective it is interesting to analyze the other two general trends in user interactions with search engines (switches and oscillations). In particular, persisting switches are interesting as they indicate a significant evolution in user preference. From a search engine's perspective it is invaluable to understand the reasons for such switches and a predictive user model that would warn of such a switch could have significant implications for the search providers involved. Similarly, an oscillating preference in search engines might represent complementary properties of engines and understanding the causal factors can help improve engine performance.

Given these key patterns of long-term search engine usage, we now seek to understand causal variables and build predictive models that can detect such behaviors.

3.2 Indicators of User Behaviors

The analyses in the previous section were based on search engine usage. We also have several other features that represent user satisfaction and query behavior: e.g., number of queries, average query length, proportions of navigational queries, etc. In this section we are interested in finding features / indicators that can distinguish between usage trends that constitute the three key behavioral trends.

To explore this we build on top of the NNMF analysis. We use the decomposition to filter and partition users that clearly exhibit the three key trends in their long-term usage. Specifically, we use the weights from the encoding matrix H to rank-order users. The weight in each row of H corresponds to the weight of the particular basis vector in each of the users. For example, we can sort by the values of the weight corresponding to the fourth basis vector (i.e., the fourth row in H) to rank order users that exhibit a persisting switch from engine B to A. Once the users are sorted we considered the first 500 users as representatives who distinctively exhibit the key trend.

We use the aforementioned method to construct three sets of 500 users who correspond to the three key trends of 1) sticking with one engine, 2) making a switch that persists, and 3) oscillating between engines. Given these sets of users, we can now analyze indicators / features that differ across groups. Statistical testing is performed using analyses of variance (ANOVA) with Tukey post-hoc testing (Q) as appropriate.

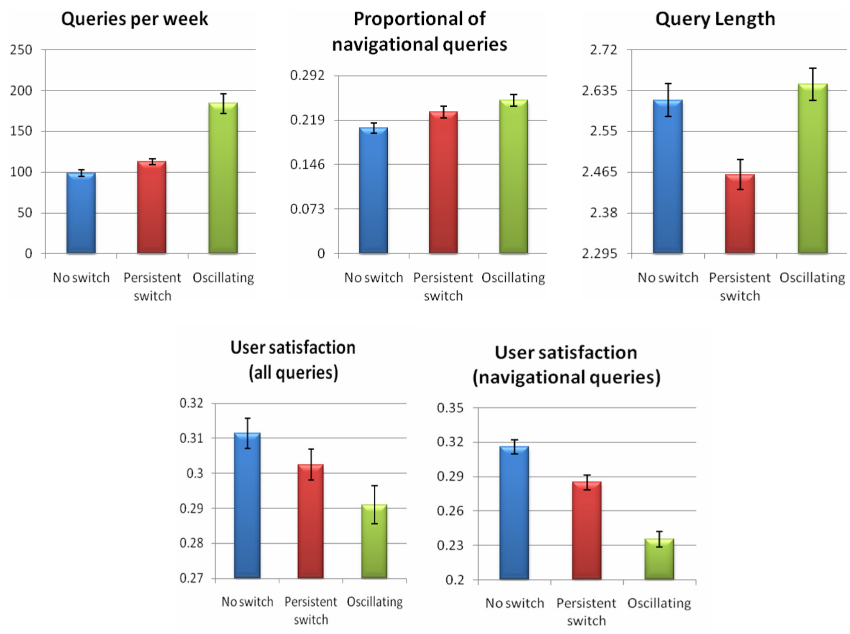

Fig. 2. Comparison of different features for the three different groups. We show means across 500 users for each group. The error bars denote the standard error.

Figure 2 shows results for different variables across the three different groups. We first observe that the users in the oscillating group issue a significantly larger number of queries than the other two groups ($F(2,1497) = 36.5$, $p < .001$; $Q \geq 6.65$, $p \leq .001$). Query frequency has previously been shown to be related to searcher expertise, and this suggests that oscillating users may be more skilled [19]. This heightened sophistication is also consistent with their awareness and usage of multiple search engines. Second, we see that these oscillating users issue a higher proportion of navigational queries than the others ($F(2,1497) = 6.2$, $p = .002$; $Q \geq 3.73$, $p \leq .02$). This may be because oscillating users are more familiar with Web resources and issue more queries to search engines requesting access to those resources. Third, we observe that oscillating users are less satisfied, both in general and for navigational queries (both $F(2,1497) \geq 4.6$, both $p \leq .01$; $Q \geq 3.44$, $p \leq .03$). These sophisticated users may pose harder queries which search engines do not perform well on or they may be more demanding in terms of what information is required to satisfy them. Finally, we notice that the query length is significantly smaller for persistent switchers ($F(2,1497) = 9.6$, $p < .001$; $Q \geq 3.56$, $p \leq .03$). If we consider the query length as a proxy for query complexity then we can hypothesize that those who are switching and sticking with the new search engine issue simple queries and might represent a population that is less familiar with search engines, a conjecture supported by [19].

Since user satisfaction has been shown to be important in brand loyalty, we examined the correlation between the frequency of usage and satisfaction. Pearson's correlation coefficients (r) between engine usage and user satisfaction for all queries were statistically significant although generally low ($-.14 \leq r \leq .36$, for the three engines), and the correlations between usage and satisfaction for navigational queries were slightly lower ($-.11 \leq r \leq .24$). Interestingly, the correlation between usage and user satisfaction varied dramatically between search engines. Engine A's usage patterns even exhibited a negative correlation with satisfaction, suggesting that people who use Engine A more often are less satisfied than those who use it less often. This engine was the most popular engine in our sample and this finding suggests that factors beyond satisfaction (e.g., brand loyalty, familiarity, or default search provider settings) may be more important in determining its usage than on the other engines.

3.3 Predicting Search Engine Usage

We aim to predict the search engine usage trend for every user from their past behavior. Specifically motivated by the discussion in the previous sections, we are interested in determining if the user will: (i) stick to their choice of search engine, (ii) switch to another engine and persist with the switch or (iii) oscillate between different search engines. Further we want to make these predictions as early as possible, since early prediction can be used by search engines to prevent unwanted switches or oscillations.

In order to explore this question, we construct a dataset consisting of 500 users from each of the three key user trends (1,500 in total). Since a majority of users belonged to the trivial *No Switch* condition, to obtain a reasonably balanced classification set we selected these 1,500 users (which was also the maximal set where all three classes were balanced). We consider examples as belonging to the *No Switch* class if the user had a single dominating search engine for more than 22 weeks in the 26 week period (i.e., same search engine for more than 85% of the weeks). Similarly, we consider users as belonging to the *Persistent Switch* class if following at least three weeks of use of a particular engine they switch to another and persist for at least eight weeks. Note that while there might be some minor oscillating characteristics in these cases, we still consider them as users belonging to persistent switch category as there was at least one switch that did persist for a significant period. The remainder of the examples are considered to belong to the *Oscillating* class. Our aim, thus, is to see if we can build a good predictive model to discriminate amongst these three classes.

The analysis from the previous section provides insights about how such predictions can be made. Specifically, Table 1 summarizes features that capture general query characteristics as well as satisfaction levels of users. Given these observations about users' past behavior we compute statistics (mean, max, min) that summarizes the usage over past weeks for all the features described in Table 1. In addition, we also compute binary features that indicate: (i) if the user has a single search engine that is dominant for more than 90% of the time observed so far (isOneEngineDominant), and (ii) if user had already made any persisting switch (observedPersistSwitch).

We performed experiments where we first split the 1,500 examples into random 50-50 train-test splits. The training data is used to learn linear classifiers with a one-vs.-all design using a Gaussian Process Regression (GPR) approach [15]. For all our experiments we used a linear kernel and set the noise parameter as $\sigma^2 = 0.1$. Once the model is trained, we can then predict class labels on the test points. Further, we are

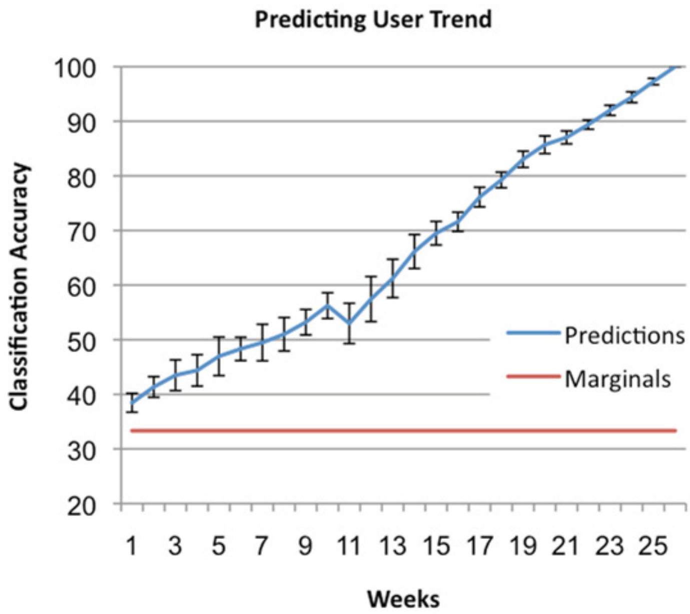

Fig. 3. Classification accuracy as we vary the number of past weeks observed. Accuracy improves as more and more past data is observed. The error bars denote standard deviation over 10 random train-test split.

also interested in determining how soon the model can predict the user trend; hence, we run these experiments for varying the number of past weeks as input. This whole methodology is run for 10 different random train-test splits. Figure 3 shows the mean recognition accuracy on the test set as we vary the number of weeks for which the data has been observed. We observe that even with one week of data, the model can predict better than random (marginal model with accuracy 33.3%). As the model observes an increasing amount of data the prediction accuracy improves constantly, eventually reaching an accuracy of 100% when all the information is observed at week 26. Note that achieving 100% accuracy at week 26 is not surprising as the test labels for these data points were originally generated by looking at 26 weeks of data. However, the experiment described above shows the promise of user modeling in predicting user trend much earlier before 26 weeks (for example even predicting based on one week's observations leads to better than chance accuracy).

Next, we analyzed the features that are helping most in classification by looking at the absolute weights in the learned linear model. A high absolute value of a weight corresponding to a feature indicates high importance in the classification. In particular, for each of the one-vs.-all classifiers the most likely prediction takes the following form [15]: $y=\mathbf{w}^T\cdot\mathbf{x}$

Here, $\mathbf{x}$ denotes the observation vector and $\mathbf{w}$ is the learned classifier. The magnitude of every i^{th} component w_i in $\mathbf{w}$ represents the contribution on the feature x_i in the classification. Consequently, we can sort the features used in prediction by the absolute

Table 2. Most discriminatory features for classifying user trends via observations to week 10

No Switch vs. All	Switch vs. All	Oscillate vs. All
isOneEngineDominant	min fractionEngine A	min fractionEngine C
min fractionEngine A	min fractionEngine C	isOneEngineDominant
observedPersistSwitch	min fractionEngine B	observedPersistSwitch
max fractionEngine A	max fractionEngine A	min fractionEngineSAT C
min fractionEngine B	max fractionEngine C	mean fractionEngineSAT A
mean fractionEngineSAT A	isOneEngineDominant	min fractionEngine B
mean fractionEngine A	max queryCountEngine C < 50	mean fractionEngineSAT B
min fractionNavEngine A	min fractionEngineSAT C	mean fractionEngineSAT C
mean fractionNavEngine A	mean fractionNavEngine A	max queryCountEngine B < 50
max fractionEngine C	observedPersistSwitch	min fractionEngineSAT B

value of the corresponding weight in the linear classifier. Table 2 shows the top 10 features selected for classification at week 10 using the average absolute value of weights across the 10 random splits. The two computed binary features (isOneEngine-Dominant and observedPersistSwitch) are important across all three classifiers. Further, it is interesting to note that these important features constitute not only the fraction of search engine usage but also the statistics about the satisfaction as well as characteristics about the queries being issued.

4 Discussion

In this work, we present evidence that there are characteristic trends in search engine usage. We also show that it is possible to develop predictive user models of these trends using a small number of features about previous engine usage and satisfaction estimates based on user interactions with search results. These findings can be used in several ways to impact the domain of search engines.

The methods general methods that we present in this paper to characterize usage patterns over time are applicable to other data sets, but the specific parameter values (e.g., the number of weeks used to help define classes of switching in our case) may need to be adapted for other data sets. We believe that our findings of a small number of consistent and predictable patterns have implications for improving the design, marketing, and user experience with search engines.

The ability to identify key trends provides insights into the different usage patterns of individual search engines as well as differences across engines. By understanding the strengths and weaknesses of different engines, appropriate resources could be directed towards general search engine improvements. In addition, the ability to predict what a searcher might do could help design adaptive search interfaces that improve relevance for particular query types, provide additional user support for query specification, or develop alternative result presentation methods.

From a marketing perspective, the ability to accurately identify different classes of users could allow search providers to target resources (e.g., marketing campaigns, incentives not to switch, etc.) to users who are the most likely to switch from (or to) their search engine. The development of richer user models could support finer-grained long-term usage analysis and more specific modeling of users' behaviors and interests. All these measures can help the search engine grow its customer base and be more sensitive to user interaction patterns and satisfaction levels.

The key method that enables the above scenarios is predictive user modeling that classifies users based on their past trends of search interactions. There is ample potential in exploring methodologies that would exploit the temporal structure of the data with richer features to provide more accurate classification as early as possible. Another possibility is to improve the classification accuracy by using a more powerful classification method, perhaps involving complex weighted committees [3].

5 Conclusions and Future Work

In this paper we have modeled aspects of multiple search engine usage over a period of six months. Our analysis identified three main classes of search engine users: those who stick with the same engine, those who switch then stick, and those who oscillate between search engines. We observed differences in how users in each of these classes interact with search engines and offered some explanations for their behavior. We also showed a small but significant correlation between search engine switching and our measure of user satisfaction. Finally, we developed a classifier to predict usage trends given historic engine usage data and satisfaction estimates, and showed that it could rapidly improve its accuracy with more data, surpassing the marginal model after only one week.

Future work involves incorporating more features into our models, such as user demographics and more detailed information about the types of queries they are issuing. This will allow us to identify additional classes of search engine usage and predict usage trends with greater levels of accuracy. Further, we are also interested in incorporating the predictive user models in order to improve search engine design and user experience.

References

1. Anderson, E.W., Sullivan, M.W.: The Antecedents and Consequences of Customer Satisfaction for Firms. Marketing Science 12, 125–143 (1993)
2. Anderson, R.E., Srinivasan, S.S.: E-Satisfaction and E-Loyalty: A Contingency Framework. Psychology and Marketing 20(3), 123–138 (2003)
3. Blum, A., Burch, C.: On-line Learning and the Metrical Task System Problem. Machine Learning 39(1), 35–58 (2000)
4. Bolton, R.N.: A Dynamic Model of the Duration of the Customer's Relationship with a Continuous Service Provider: The Role of Satisfaction. Marketing Science 17(1), 45–65 (1998)
5. Capraro, A.J., Broniarczyk, S., Srivastava, R.K.: Factors Influencing the Likelihood of Customer Defection: The Role of Consumer Knowledge. Journal of the Academy of Marketing Science 31(2), 164–175 (2003)

6. Fox, S., Karnawat, K., Mydland, M., Dumais, S.T., White, T.: Evaluating Implicit Measures to Improve the Search Experience. ACM Transactions on Information Systems 23(2), 147–168 (2005)
7. Heath, A.P., White, R.W.: Defection Detection: Predicting Search Engine Switching. In: Ma, W.Y., Tomkins, A., Zhang, X. (eds.) World Wide Web Conference, pp. 1173–1174 (2008)
8. Juan, Y.F., Chang, C.C.: An Analysis of Search Engine Switching Behavior Using Click Streams. In: Deng, X., Ye, Y. (eds.) WINE 2005. LNCS, vol. 3828, pp. 806–815. Springer, Heidelberg (2005)
9. Keaveney, S.M., Parthasarathy, M.: Customer Switching Behavior in Online Services: An Exploratory Study of the Role of Selected Attitudinal, Behavioral, and Demographic Factors. Journal of the Academy of Marketing Science 29(4), 374–390 (2001)
10. Laxman, S., Tankasali, V., White, R.W.: Stream Prediction Using a Generative Model Based on Frequent Episodes in Event Sequences. In: Li, Y., Liu, B., Sarawagi, S. (eds.) ACM SIGKDD Conference on Knowledge Discovery and Data Mining, pp. 453–461 (2008)
11. Lee, D.D., Seung, S.: Learning the Parts of Objects by Non-negative Matrix Factorization. Nature 401, 788–791 (1999)
12. Mittal, B., Lassar, W.M.: Why do Customers Switch? The Dynamics of Satisfaction versus Loyalty. Journal of Services Marketing 12(3), 177–194 (1998)
13. Mukhopadhyay, T., Rajan, U., Telang, R.: Competition Between Internet Search Engines. In: Hawaii International Conference on System Sciences (2004)
14. Pew Internet and American Life Project: Search Engine Users (2005) (accessed December 2008)
15. Rasmussen, C.E., Williams, C.K.I.: Gaussian Processes for Machine Learning. MIT Press, Cambridge (2006)
16. Sasser, T.O., Jones, W.E.: Why Satisfied Customers Defect. Harvard Business Review (November-December 1995)
17. Telang, R., Mukhopadhyay, T., Wilcox, R.: An Empirical Analysis of the Antecedents of Internet Search Engine Choice. In: Workshop on Information Systems and Economics (1999)
18. White, R.W., Dumais, S.T.: Characterizing and Predicting Search Engine Switching Behavior. In: Cheung, D., Song II, -Y., Chu, W., Hu, X., Lin, J., Li, J., Peng, Z. (eds.) ACM CIKM Conference on Information and Knowledge Management, pp. 87–96 (2009)
19. White, R.W., Morris, D.: Investigating the Querying and Browsing Behavior of Advanced Search Engine Users. In: Clarke, C.L.A., Fuhr, N., Kando, N., Kraaij, W., de Vries, A.P. (eds.) ACM SIGIR Conference on Research and Development in Information Retrieval, pp. 255–261 (2007)
20. White, R.W., Richardson, M., Bilenko, M., Heath, A.P.: Enhancing Web Search by Promoting Multiple Search Engine Use. In: Myaeng, S.-H., Oard, D.W., Sebastiani, F., Chua, T.S., Leong, M.K. (eds.) ACM SIGIR Conference on Research and Development in Information Retrieval, pp. 43–50 (2008)