# Taking into account Tabbed Browsing in Predictive Web Usage Mining

Geoffray Bonnin, Armelle Brun and Anne Boyer
*LORIA – KIWI Team*
*Campus Scientifique – BP 239*
*54506 Vandoeuvre-lès-Nancy Cedex, France*
{*geoffray.bonnin, armelle.brun, anne.boyer*}*@loria.fr*

***Abstract*—Over the last few years, browser tabs have become a very common tool for web users and have been extensively used to perform parallel navigations. Tabbing facilitates web browsing but results in an imbrication of navigations, which makes it more difficult to understand users' behavior. That is why very recent research has been focusing in analyzing this new kind of usage. This work follows a previous publication in which a new model was proposed to model parallel browsing. In this paper, we propose a new strategy to better take into account tabbing activity. Experiments are performed on an open browsing dataset. Results show that our model provides an accuracy similar to the one of a state-of-the-art model that implicitly takes into account parallel browsing. It thus constitutes a strong basis to estimate tabbing activity. We then present the statistics about parallel browsing that our approach provides.**

***Keywords-web usage mining; web predictive modeling; Markov models; tabbing***

## I. INTRODUCTION

Parallel browsing consists in performing several web navigations at the same time by successively switching from one to another. Until the late 90's, this activity could only be performed by using several browser windows. At that time, studies showed that multiple browsing windows were used less than $1\%$ of the time [9], [27]. Since 2000, when the Opera browser first introduced the *tab* mechanism, it has been possible to perform parallel browsing within one single window. From that time, this feature has progressively been offered by every web browser. Numerous recent studies have shown how much this simple interface changed user behavior, and how much current web users extensively perform parallel browsing [15], [18], [29], [30]. The same conclusion can also be drawn from the Test Pilot data provided in April 2010 by Mozilla Labs [2]. Indeed, we performed a quick analysis of this data and found that every single user had performed tabbing, and that more than half of the users had used more that ten tabs during one session.

It is thus clear that parallel browsing has now become a very common activity. Such an information means that traditional web usage mining approaches [19], [25] can no longer be used in their linear form and that new strategies have to be studied. One of the major application of web usage mining is predictive modeling. Indeed, predictive modeling is useful for many purposes such as web page research [26], latency reduction [23], arrangement of the links in a website [11], web recommendation [20], etc. The most popular techniques used in this domain are association rules [3], sequential patterns [4] and Markov models [22], none of which takes into account parallel browsing. The usual challenge of predictive modeling is to provide a model that is a trade-off between predictive accuracy, coverage, and space and time complexity [5], [7], [13], [22].

When a user performs several parallel navigations using several tabs or several windows, then the resulting session recorded in the logs is an imbrication of linear navigations that cannot directly be identified. This phenomenon is illustrated in Figure 1: a user performs two parallel navigations $\langle A_1, A_2, A_3\rangle$ and $\langle B_1, B_2, B_3\rangle$ using two tabs, which induces the mixed session $\langle A_1, A_2, B_1, B_2, A_3, B_3\rangle$ to be recorded in the logs. Although it is possible to avoid this phenomenon by including tabbing information directly within the logs, as done for instance in the Mozilla Labs Test Pilot data, usual logs do not contain this kind of information. Thus, dealing with the retrieval of linear navigations from such raw logs seems a very promising way to enhance the accuracy of web predictive modeling.

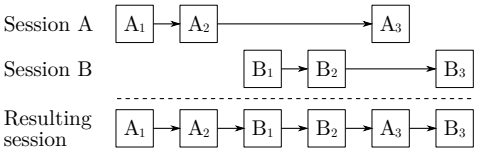


Figure 1. Imbrication of two linear navigations through parallel browsing

In a previous work, we proposed a new model for tabbed browsing called Tabbing-Based All-$k^{\text{th}}$-Order Markov model (TABAKO) [6]. In this paper, we propose a new strategy to extract linear sessions that puts away the concept of tasks. We show that the resulting version of the TABAKO model provides an accuracy comparable to the state-of-the-art and as it explicitely models parallel browsing, it can be used to accurately measure tabbing activity.

The rest of this paper is organized as follows. We are first interested in works related to parallel browsing. We then present our new proposition to model parallel browsing in Section III. Experimental setup is specified in Section IV and the results of our experiments are discussed in Section V.

 ISBN: 978-1-61208-163-2 

Conclusion and perspectives are put forward in the last section.

## II. Related Work

In this section, we first present previous works in the field of predictive modeling that take into account parallel browsing. We then turn to the only work, to best of our knowledge, in which tabbing activity is estimated from raw logs.

### A. Predictive Modeling for Parallel Browsing

Before the existence of the tabbing mechanism in recent browsers, several approaches had been proposed to discover and refine sessions in web usage data, some of which can be compared to our purpose. For instance, Chen et al. [10] proposed to use a concept called *Maximal Forward Reference* (MFR), *i.e.*, the minimal path from a start page to a target page. More precisely, a MFR is obtained by filtering backward references that are induced by the use of the back button. The major difference with such approaches is that they do not allow to take into account the imbrications induced by the use multiple tabs or multiple windows.

More recently, some works have dealt with approaches that are able to implicitly take into account parallel browsing. In this scope, Jin et al. [17] proposed a web recommendation system able to discover task-level patterns. The authors use probabilistic latent semantic analysis to characterize users' navigational tasks. They then use a Bayesian updating to compute the probability of each task being performed according to a given active session. Then recommendations are computed using a maximum entropy model in which one of the features uses a first-order Markov model. In the same spirit, we proposed in [5] to implicitly take into account parallel browsing by using a skipping-based model. This model is based on a non-contiguous Markov model and computes recommendations based on a weighted combination of subhistories, *i.e.*, small subsequences of user's current session.

In [12], Chierichetti et al. propose to use a simple branching process that captures the opening, switching and closing of tabs to enhance a simple Markov model. The distances from true distributions of this model and PageRank [8] are then compared. Results show that taking in to account tabbing enhances the accuracy of the modeling. However, the data used by the authors contains information about the source node from which users arrive to each page, which allows to deduce information about tabbing activity. Such an information is not commonly available. Also, handling such a branching process induces a large time and space complexity and is not appropriate for common web usage mining applications.

In this paper, we thus focus on the ability to explicitly model parallel browsing and imbrications of navigations without any specific information about tabbing activity.

### B. Tabbing Activity Measure

To the best of our knowledge, the work of Viermetz et al. (2006) [28] is the only work in which tabbing activity is estimated from logs with no explicit information about tabbing activity. Such an estimation is obtained by building what the authors call a *clicktree model*, *i.e.*, a tree into which all the possible tabbed paths of user logs are stored. Using this model, the authors estimated that users perform tabbing from 4% to 85% of the time, which is a quite large range. Moreover, as for the work of Chierichetti et al. [12], handling such clicktrees involves huge time and space complexities and is not appropriate for common web usage mining applications.

As will be shown in this paper, the model we propose provides an accurate estimation of tabbing activity while having a low time and space complexity.

## III. A New Strategy for the Extraction of Linear Sessions

As mentioned in the introduction, we proposed in [6] a first model for parallel browsing we called Tabbing-Based All-$k^{\text{th}}$-Order Markov model (TABAKO). We first briefly recall the principles on which it is based and then present our new strategy to extract the linear sessions. The rest of the functioning of the TABAKO model remains similar.

### A. Overview of the Previous Work

Our previous work was based on a concept we called *tasks* and defined as being a typical sequence of resources that can have several slight variations. We further defined this concept through a global alignment algorithm, *i.e.*, the aforementioned variations correspond to insertions, deletions and substitutions in the sequence of resources. More specifically, we considered that two sessions correspond to a similar task if the score of their best global alignment computed using the Needleman-Wunsch algorithm [21] was greater than a predefined threshold $t_1$. In the same spirit, we also considered that a session $X$ is a subsession of $Y$ if the score of their best local alignment using a modified version of the Smith-Waterman algorithm [24] was greater than a given threshold $t_2$. Thus, the basis idea was to say that if a session contains two different tasks, then this session is not a linear session.

The general functioning of the TABAKO model is the following. First, the linear sessions of the training corpus are first extracted using the aforementioned corpus and then stored in order to be used during the prediction step. Second, the model is built exactly as an all-$k^{\text{th}}$-order Markov model [22], except that only the linear sessions are used to train the model. Then, predictions are computed in two steps: (1) the extraction of the best overlapping and the extraction of the corresponding linear sessions, and (2) the creation of the prediction lists by applying the all-$k^{\text{th}}$-order Markov model on the retrieved linear subsessions.

 ISBN: 978-1-61208-163-2

### B. Improvement of the Extraction of Linear Sessions

We focus here on a new strategy for the extraction of the linear sessions within logs. It is based on the following considerations. First, the concept of tasks we defined in our previous work may conflate several different concepts that, while related, are not necessarily comparable. In particular, it is possible that a user performs one single task using several tabs. Also, there is no evidence that the notion of sequence alignment we used does correspond to actual tasks performed by users. Last, such a strategy involves the computation of many sequence alignments and induces a high time complexity.

The strategy we propose here consists in putting away the concept of tasks and thus putting away sequence alignment algorithms. Two important changes are then induced. The first is the replacement of the local alignment algorithm used for the identification of subsessions by a direct matching algorithm. As will be discussed in the following, this reduces the time complexity of the algorithm. The second consists in not performing the task discrimination step, which allows to take into account the situation when a user opens several tabs to perform one single task. Thus all previous considerations are taken into account, and the integration of the resulting algorithm into the TABAKO model should enhance its accuracy.

The general functioning of the extraction of linear sessions can then be summarized as follows: given a session $s$, we first check whether a session $\ell_1$ in the corpus is a subsession of $s$, *i.e.*, each element of $\ell_1$ can be found in $s$ in the same order. If such a session is found, we remove the corresponding elements in $s$, which results in a new subsession $r$. We then check whether a session $\ell_2$ in the corpus is a subsession of $r$. If such a session is found, then $s$ in not a linear session. The corresponding algorithm is detailed in Algorithm 1. The idea behind the algorithm is that some of the sessions recorded in the logs correspond to linear navigations, *i.e.*, they are performed using a single window and a single tab. Thus, these sessions can be used to detect the imbrications within the non-linear navigations.

The rest of the functioning of the TABAKO model remains the same, except that subsessions are identified through an exact matching algorithm.

### C. Time Complexity

We now focus on the time complexity of the algorithms. We first consider the time complexity of the algorithms used for the identification of subsessions. Recall that the purpose of this process is to determine whether a session $\ell$ is a subsession of a session $s$. Using a direct matching algorithm, the time complexity is $\mathcal{O}(|\ell|+|s|)$. Using a sequence alignment algorithm, the time complexity is $\mathcal{O}(|\ell| \cdot |s|)$. Considering that the sessions of the corpus have a maximum size of $M$, the respective complexities are $\mathcal{O}(M)$ and $\mathcal{O}(M^2)$. Thus, using a direct matching algorithm for the identification of subsessions induces a very smaller time complexity.

**Algorithm 1** Extraction of the linear sessions

```
Input: A list of sessions S
Output: The corresponding list of linear sessions
  for all session s in S do
    for all session ℓ1 ≠ s in S do
      if isSubsession(ℓ1, s) then
        r ← s − ℓ1
        for all session ℓ2 in S do
          if isSubsession(ℓ2, r) then
            remove s from S
          end if
        end for
      end if
    end for
  end for
  return S
```

We now consider the time complexity of the general algorithm for the extraction of linear sessions. Using our new proposition, each session $s$ of the corpus is compared to all the other sessions until a matching subsession $\ell_1$ is found using a direct matching algorithm. Thus, for each session $s$, if the training corpus contains $N$ sessions, the corresponding time complexity is $\mathcal{O}(N \cdot M)$. Once such a subsession $\ell_1$ has been found, the matching elements are removed from $s$, and another subsession $\ell_2$ is searched in the training corpus, which is also perform in $\mathcal{O}(N \cdot M)$. Thus, the time complexity of the general algorithm is $\mathcal{O}\left(N \cdot (N \cdot M + N \cdot M)\right) = \mathcal{O}(N^2 \cdot M)$. Using our previous proposition, each session $s$ of the corpus is compared to all the other sessions to find all possible subsessions using a local alignment algorithm, and then all such subsessions are compared two by two using a global alignment algorithm until two of them correspond to different tasks. The corresponding time complexity is $\mathcal{O}(N^3 \cdot M^2)$. Thus, our new strategy has a very lower time complexity.

## IV. Experimental Setup

### A. Evaluation Metric

In order to evaluate the accuracy of our model when our new strategy is integrated, we use the *hit ratio* [14], [16], [22]. For each session of the test corpus and for each browsing step, a prediction list of size $m$ is built, containing the most probable resources according to the model. A hit means that the resource the active user has actually consulted is in the list. In the following experiments we use lists of size 10.

### B. Data

Empirical studies are performed on the CTI web server corpus of the DePaul University [1]. It contains 69,471

 ISBN: 978-1-61208-163-2

consultations of 683 pages by 5,446 users during a two-weeks period in April 2002 (*i.e.*, about 170 consultations per day). The data provided has been cleaned and filtered by eliminating sessions of size 1 and low support page views.

When a session starts, the highest order Markov model that can be used is a Markov model of order 0. As well, after a user has browsed one resource, the highest order Markov model that can be used is a Markov model of order 1. Thus, the differences in predictive ability between more sophisticated models only appear beyond this scope and sessions of size 1 and 2 are not interesting for this purpose. That is why we also eliminated sessions of size 2 from the corpus.

The repartition of session sizes of the resulting corpus is depicted in Figure 2. As it can be seen, most of the sessions ($66.0\%$) have a size between 3 and 5. The corpus has an average session size of $5.9$ and a standard deviation of $4.1$. Assuming that small sessions are more likely to be linear, we can guess that the proportion of linear session is more than $66.0\%$.

The corpus has been divided into a training and a test set of $90\%$ and $10\%$ respectively.

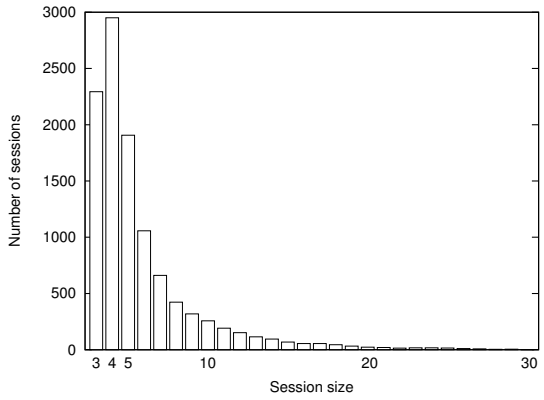


Figure 2. Session sizes of the DePaul university browsing dataset

## V. Experimental Results

This section is dedicated to the results of our algorithms. We introduced a new strategy to extract linear sessions from logs, which is one step of the functioning of the TABAKO model. We first study the impact of this change on the accuracy provided by the model. In this frame, we compare the model to the previous proposition and to the state-of-the-art. We then put forward the ability of our model to measure tabbing activity and present the statistics about parallel browsing it provides.

### A. Comparison to the State-of-the-art

We are first interested in the accuracy of the TABAKO model when the concept of tasks is put away. We also compare our model to one of the best performing models of the state-of-the-art, the SBR model. This model implicitly takes parallel navigations into account and have already proved its high modeling accuracy [5].

Hit ratios of the old version (with tasks) and of the new version (without tasks) of the TABAKO model, and of the SBR model are presented in Table I. First, we can notice that compared to the previous version of the TABAKO model, the accuracy is enhanced by $12\%$. This confirms that the concept of tasks we used in our previous proposition is not suited for modeling parallel browsing and puts forward the better relevance of our new strategy.

The most interesting result is that the new version of the TABAKO model provides results similar to the ones of the SBR model. The fact that the model provides similar results is interesting because it explicitly takes into account parallel browsing. This means that this new version of the TABAKO model now constitutes a strong basis to estimate statistics about tabbing activity.

Table I
Comparison to the state-of-the-art

| | TABAKO (no tasks) | TABAKO (tasks) | SBR |
|---|---|---|---|
| Hit ratio | **60.1** | 54.7 | 60.1 |

### B. Estimation of the Tabbing Activity

We are now interested in the ability of our algorithms to estimate tabbing activity. Recall that one of the major difference between models like the TABAKO model or the SBR model on one hand, and previous contributions of the state-of-the-art that tried to refine sessions in web usage data on the other hand, is the ability to model imbrications of navigations. However, so far, no information guarantees that such imbrications are contained within the dataset on which we experimented. Thus, we are not only interested in estimating the proportion of non linear sessions, but also the proportion of sessions that contain imbrications. The corresponding results are presented in Table II.

As it can be seen, very few parallel navigations are detected: only 11.7% of the sessions are considered as being non linear, although recent studies show that most of the navigations are not linear [15], [18], [29], [30]. However, this value is plausible for three reasons: first, the data dates back to 2002, and tabbing may not have been much used at that time. Second, studies indicating more than 50% of parallel browsing are based on inter-sites navigation, whereas this corpus contains intra-site navigations, which are less likely to involve parallel browsing. Last, as it can be seen in Figure 2 of Section IV, most of the sessions in this corpus are rather small, and thus may not contain parallel browsing. This last consideration is confirmed by the fact that the

 ISBN: 978-1-61208-163-2 

Table II
STATISTICS ABOUT TABBING ACTIVITY

| | |
|---|---|
| Proportion of linear sessions | 88.3% |
| Average size of linear sessions | 5.0 |
| Standard deviation for linear sessions | 2.7 |
| Proportion of non linear sessions | 11.7% |
| Average size of non linear sessions | 12.2 |
| Standard deviation for non linear sessions | 6.7 |
| Proportion of imbricated sessions | 2.0% |

extracted linear sessions have an average size of $5.0$, while the extracted non linear sessions have an average size of $12.2$.

Focusing on the imbrications, we can see that $2.0\%$ of the session do contain such imbrications. As $11.7\%$ of the sessions are not linear, this means that when a user uses several tabs, the corresponding sessions are imbricated linear navigations $17\%$ of the time. This argues in favor that although dating back from 2002, the dataset on which we performed our experiments did contain parallel browsing and that it is necessary to use algorithms able to take into account this phenomenon in order to build accurate predictive models. As parallel browsing was still a new usage in 2002, this proportion should be higher in more recent data. Thus, the difference in accuracy with models that do not take into account imbrications should be higher.

Generally speaking, the statistics provided by the TABAKO model seem likely to correspond to the reality of this particular data, and are further consolidated by the predictive accuracy of the model. Thus, although this model seems to constitute a reasonable tool for measuring tabbing activity, it should be further experimented on a newer dataset.

## VI. CONCLUSION AND FUTURE WORK

In this article, we focused on predictive modeling for parallel browsing. We proposed a new strategy for the extraction of linear sessions from logs in which no information about parallel browsing is provided. We then integrated this new strategy into a previous proposition called the Tabbing-Based All-$k^{\text{th}}$-Order Markov model (TABAKO). The new version of the model has a drastically lower time complexity and allows to take into account the case when a user performs one task using several tabs.

Our model was studied on an open browsing dataset, and provided an accuracy similar to one of the best performing state-of-the-art predictive model. We then concluded that it constitutes a strong basis to estimate statistics about tabbing activity. These statistics show that very few parallel navigations are contained in the corpus, which is plausible according to its age. Moreover, the statistics show that imbrications are contained in the navigations, which argues in favor of the necessity to use algorithms able to take into account this phenomenon in order to build accurate models.

In a future work, we plan to enhance the predictive accuracy of the model by investigating two new strategies. The first one is to propose a new incremental algorithm for the extraction of linear sessions, by first considering the smallest sessions as being linear and then moving to larger sessions. The second is to propose a new algorithm that uses a tree structure to determine the best imbrication of linear sessions in each non-linear sessions. We also plan to incorporate other elements in the modeling of web navigation, such as the use of the back button, which was not explicitly taken into account in this work.

## REFERENCES

[1] Preprocessed DePaul CTI Web Usage Data. Retrieved September 19, 2011. http://maya.cs.depaul.edu/classes/ect584/resource.html, 2002.

[2] Tab Switch Study: Aggregated Data Samples. Retrieved September 19, 2011. https://testpilot.mozillalabs.com/tabswitch/aggregated-data.html, 2010.

[3] R. Agrawal, T. Imielinski, and A. Swami. Mining Association Rules between Sets of Items in Large Databases. In P. Buneman and S. Jajodia, editors, *Proceedings of the 1993 ACM SIGMOD International Conference on Management of Data*, pages 207–216, 1993.

[4] R. Agrawal and R. Srikant. Mining Sequential Patterns. In *ICDE'95: Proceedings of the International Conference on Data Engineering*, pages 3–14, 1995.

[5] G. Bonnin, A. Brun, and A. Boyer. *Skipping-Based Collaborative Recommendations inspired from Statistical Language Modeling*, chapter 13, pages 263–288. Web Intelligence and Intelligent Agents, INTECH, 2010.

[6] G. Bonnin, A. Brun, and A. Boyer. Towards Tabbing Aware Recommendations. In *Proceedings of the First International Conference on Intelligent Interactive Technologies and Multimedia*, pages 316–323. ACM, 2010.

[7] J. Borges and M. Levene. Generating dynamic higher-order markov models in web usage mining. 2005.

[8] S. Brin and L. Page. The Anatomy of a Large-Scale Hypertextual Web Search Engine. *Computer networks and ISDN systems*, 30(1-7):107–117, 1998.

[9] L. Catledge and J. Pitkow. Characterizing Browsing Strategies in the World-Wide Web. *Comput. Netw. ISDN Syst.*, 27(6):1065–1073, 1995.

 ISBN: 978-1-61208-163-2 

[10] M. Chen, J. Park, and P. Yu. Data Mining for Path Traversal Patterns in a Web Environment. In *Proceedings of the 16th International Conference on Distributed Computing Systems*, pages 385–392. IEEE Computer Society, 1996.

[11] E. Chi, J. Pitkow, J. Mackinlay, P. Pirolli, R. Gossweiler, and S. Card. Visualizing the Evolution of Web Ecologies. In *CHI '98: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 400–407, 1998.

[12] F. Chierichetti, R. Kumar, and A. Tomkins. Stochastic Models for Tabbed Browsing. In *Proceedings of the 19th international conference on World wide web*, pages 241–250. ACM, 2010.

[13] M. Deshpande and G. Karypis. Selective Markov Models for Predicting Web Page Accesses. *ACM Trans. Internet Technol.*, 4(2):163–184, 2004.

[14] S. Gündüz and M. Özsu. A Web Page Prediction Model Based on Click-Stream Tree Representation of User Behavior. In *KDD '03: Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 535–540, 2003.

[15] J. Huang and R. White. Parallel Browsing Behavior on the Web. In *Proceedings of the 21st ACM conference on Hypertext and hypermedia*, pages 13–18. ACM, 2010.

[16] X. Jin, B. Mobasher, and Y. Zhou. A Web Recommendation System Based on Maximum Entropy. In *Proceedings of the International Conference on Information Theory: Coding and Computing*, pages 213–218, 2005.

[17] X. Jin, Y. Zhou, and B. Mobasher. Task-oriented Web User Modeling for Recommendation. In *Proceedings of the 10th International Conference on User Modeling (UM)*, pages 109–118, 2005.

[18] M. Meiss, J. Duncan, B. Gonçalves, J. Ramasco, and F. Menczer. What's in a Session: Tracking Individual Behavior on the Web. In *Proceedings of the 20th ACM conference on Hypertext and hypermedia*, pages 173–182. ACM, 2009.

[19] B. Mobasher. *Data Mining for Web Personalization*, chapter 3, pages 90–135. LNCS 4321 - Brusilovsky, P. and Kobsa, A. and Nejdl, W., 2007.

[20] M. Nakagawa and B. Mobasher. Impact of Site Characteristics on Recommendation Models Based On Association Rules and Sequental Patterns. In *Intelligent Techniques for Web Personalization*, 2003.

[21] S. Needleman and C. Wunsch. A general method applicable to the search for similarities in the amino acid sequence of two proteins. *Journal of molecular biology*, 48(3):443–453, March 1970.

[22] J. Pitkow and P. Pirolli. Mining Longest Repeating Subsequences to Predict World Wide Web Surfing. In *USITS'99: Proceedings of the 2nd conference on USENIX Symposium on Internet Technologies and Systems*, pages 139–150, 1999.

[23] S. Schechter, M. Krishnan, and M. Smith. Using Path Profiles to Predict HTTP Requests. *Computer Networks and ISDN Systems*, 30(1-7):457–467, 1998.

[24] T. Smith and M. Waterman. Identification Of Common Molecular Subsequences, 1981.

[25] J. Srivastava, R. Cooley, M. Deshpande, and P. Tan. Web Usage Mining: Discovery and Applications of Usage Patterns from Web Data. *SIGKDD Explorations Newsletter*, 1(2):12–23, 2000.

[26] P. Tan and V. Kumar. Discovery of Web Robot Sessions Based on their Navigational Patterns. *Data Mining Knowledge Discovery*, 6(1):9–35, 2002.

[27] L. Tauscher and S. Greenberg. How People Revisit Web Pages: Empirical Findings and Implications for the Design of History Systems. *International Journal of Human Computer Studies*, 47:97–137, 1997.

[28] M. Viermetz, C. Stolz, V. Gedov, and M. Skubacz. Relevance and Impact of Tabbed Browsing Behavior on Web Usage Mining. In *Proceedings of the 2006 IEEE/WIC/ACM International Conference on Web Intelligence (WI)*, pages 262–269, 2006.

[29] H. Weinreich, H. Obendorf, E. Herder, and M. Mayer. Not Quite the Average: An Empirical Study of Web Use. *ACM Transactions on the Web (TWEB)*, 2(1):5, 2008.

[30] H. Zhang and S. Zhao. Measuring Web Page Revisitation in Tabbed Browsing. In *Proceedings of the 2011 annual conference on Human factors in computing systems*, CHI '11, pages 1831–1834. ACM, 2011.

 ISBN: 978-1-61208-163-2 