Detecting session boundaries from Web user logs

Daqing He
School of Computer and Mathematical Sciences, The Robert Gordon University
Aberdeen AB25 1HG, Scotland
dqh@scms.rgu.ac.uk

Ayşe Göker
School of Computer and Mathematical Sciences, The Robert Gordon University
Aberdeen AB25 1HG, Scotland
asga@scms.rgu.ac.uk

Abstract

Detecting session boundaries on the Web is important for several reasons. Firstly, it is important to establish a common context for various statistics relating to user sessions and frequency of user activities. More specifically, it is important to detect some boundaries in order to group related information together for other applications, such as learning techniques for adaptive search engines.

To date, however, the notion of a session on the Web has not been consistently defined, if it at all. The tendency has been to group the log data that has been made available from one user or IP address under the umbrella of one session regardless of the length of time covered by the logs. This tendency lacks a more user oriented view. Our argument is that a session on the Web can be defined as a group of user activities that are related to each other not only through an evolving information need but also through close proximity in time. Thus, we describe and discuss the investigation based on two Web transaction logs (Excite and Altavista) with a view to structuring the activities into sessions or units for subsequent use in user-oriented learning techniques.

The paper describes the methodology and the experiments performed followed by results and discussions. The results point to a 10-15 minute threshold between user activities for an appropriate session interval. The implications and limitations of the results as well as differences with traditional IR systems are also discussed.

1 Introduction and Motivation

When discussing the information retrieval process, often the focus is on the individual stages such as formulating queries, searching the document collections, and presenting the set of returned documents. However, there are cases where we need to go beyond the boundaries of the individual stages and consider the cycles of these stages. Research [8] shows that nearly 60% of users had conducted more than one IR search for the same information problem. In the research referred to, the process of repeatedly searching over time in relation to a specific, but possibly evolving information problem is defined as the **successive search phenomenon**.

The study described in this paper investigates a subset of such a phenomenon, that is groups of sequences of *activities* are related to each other not only through an evolving information need at a deeper, conceptual level but also through close proximity in time. We intend to group these activities together and refer them as a **session**. By **activities**, we refer to the search related actions which take place during the course of information retrieval. These include forming or reforming a query (more precisely a search statement), browsing the results, providing relevance judgements, and so on.[1]

A session is quite well defined (at least at the 'superficial' level) in the context of a user and a human intermediary in a library. Typically, it starts from the beginning of the dialogue between the user and the intermediary, and ends

[1] We think it is a little misleading to refer to all activities as queries. This creates a confusion between the initial search statement and any following browsing activities, for example. In this paper, only forming and modifying a search statement is referred to as a query.

when the dialogue ends. The dialogue may end for a variety of reasons. These might be that: the user finds some documents or pointers to documents he/she is looking for; the user receives some feedback as to why the information can not be found; or there is a time constraint involved. A session also has a rather clear meaning in an online search system with user login and logout facilities. A session in this case starts from the time when a user performs login, and finishes upon logout.

However, it is different on the Web. With the rapid expansion of the Internet, a great variety of information retrieval tasks are performed on the Web using search engines. Because a Web search does not usually involve a human intermediary, and the fact that there is usually no official login and logout to access and use the search engines, it is not very clear exactly what a 'unit'/session should be (at least from a system-oriented definition perspective).

Studying users in interactive systems is difficult [7], but studying users on the Web is even more difficult. Unless a special browser is used, the information available from the retrieval process about a user and his information need is much less for the Web search than a traditional library search. The queries are short [5], and with users potentially accessing the search engine from all over the world at any point in time with a variety of methods, it is even difficult to identify the user after each activity. Therefore, it is better to collect the information from a group of activities (such as those within a session), rather than from one query. However, as one of the most important preconditions, the session has to be identified. This is the aim of this investigation.

Previous research found that frequent users in a bibliographic information retrieval system usually have 2 - 3 subject matters[2] that they search around [3]. The Web users might have a different number of subject matters/topics, but it is not unreasonable to assume that their searches will tend be clustered around their interests or requirements (consciously or unconsciously). If we view a user with an interest in a particular topic as acting in a particular **role**, we could anticipate that user activities in the same session are more likely to correspond to this role. This claim could be reasonable, particularly if the retrieval process can be viewed as an interactive problem solving task with a particular goal.

Whether there are connections between roles and contexts of different user information needs is an open question, but it is reasonable to think that there are contextual connections among different searches from the same user acting the same role. The contextual information of one search can be helpful for subsequent searches as long as the role behind them are the same.

In summary, the benefits of having sessions identified for Web searches are (see section 3 for further description on sessions) :

- Web queries tend to be short, have a simple structure and are generally not based on any help from a human intermediary. It is probably difficult to identify the role for each query. However, if we can know or ensure that a session (a particular unit) only corresponds to one role, the role identification to a particular query could be helped from other queries in the same session.

- Sometimes, it is useful to know which queries are from the same user. By knowing that certain queries are in the same Web session, we can more safely presume that they belong to the same user.

- Sessions can also be used in other types of Web user studies, such as modelling and learning about Web surfing patterns, and thus a common framework for sessions based on the statistics would be useful.

The remaining parts of this paper describe and discuss the investigations on the Web Transaction Logs (TLs) with a view to structuring the transactions into sessions or units for subsequent use in user-oriented learning techniques. Firstly, the related work is described, then the methodology of our study and the experiments performed in our study are presented. This is followed by further work and conclusions.

2 Related work

Studies on Web navigation activities based on logs spanning long periods of time (e.g. weeks or more) indicate that it is very likely that users will visit a web site more than once [1, 2]. They use the idea of what they call *timeout* to divide

[2] Hjorland discusses subject matters from several perspectives [4]. This work is based on the pragmatic view he refers to.

the page accesses of each user into individual sessions [2]. A timeout is the time between two adjacent activities. Catledge and Pitkow found a 25.5 minute timeout based on their user experiments in 1994 [1]. However, the reasons behind choosing that particular amount of timeout are not clear, and the users' navigation patterns may have changed over the last six years. More importantly, Catledge and Pitkow's work is about users' navigation behaviour, and does not include activities of using Web search engines.

In Web TL Analysis on users' retrieval activities, explicitly or implicitly, studies often group all activities for one user (if registered) or IP number into a unit referred to as a *session* [5, 6]. The time period these logs cover vary from a few hours in a day [5] to several days/months [6]. The appropriateness of grouping these activities under one session is debatable, particularly where the time span is large. Additionally, one could argue that the final cut-off point for the TLs is usually arbitrary and the TLs could just as well be split into several different batches.

3 Methodology

At this point, we would like to clarify that although we have been referring to the term *session*, actually we are more concerned with detecting a 'meaningful' unit of activities (or boundaries between groups of activities) with the intention of using them for a learning technique. With this in mind, the start and end of a session are the points where the role behind a query changes. The identification of sessions themselves would be straightforward if the role behind each query/activity was known. However, automatically identifying a role is a difficult task, and requires further information about the background to the query from the session. Hence, the identification of a session has to be done through other information.

Activities in the same session are not only more likely to share the same role at a conceptual level, but also are close to each other in terms of generation time. Although there is a time gap between two adjacent activities in the same session, we think that the gap would usually be smaller than that between two activities in the different sessions. A time span called **session interval** could be defined in advance to be used as a threshold. Two adjacent activities are counted in two different sessions if the time between them exceeds this threshold. Hence, the identification of session boundaries or delimiters now effectively becomes a process of examining the time gap between activities and their number of occurrences and comparing with the session interval.

Ideally, a session should contain only those activities from and only from one role. In this respect, the **optimal session interval** should not be too large since the larger the gap, the higher the risk of grouping activities from different roles together – which results in the correctness of the role information from other queries being reduced. However, having too small a session interval also has its problems as essentially there is less information available on the role at a particular point in time.

The questions we aim to address in our approach to sessions on the Web are:

1. "*Is there a session interval that can be used for detecting sessions on the Web?*"
2. "*What is the optimal value of the session interval with regard to the likehood of belonging to one role?*"

The above two questions probably could be answered through cognitive analysis or some other conceptual model of Web users' behaviour. However, we have used a more direct empirical method, that is to analyse transaction logs from tens of thousands of user search activities. Our aim is not to find a perfect session interval that is correct in all the situations, but an *acceptable* session interval that works for *most* sessions. Our approach focused on the second question above directly. The answer for the first question is included in the answer of the second one.

The experiments presented in this paper belong to the first stage of our work on this topic. Experiments on session cut errors aiming for further analysis of session intervals are parts of subsequent work.

4 The data

The experiments were based on two sets of web logs. These were Excite (http://www.excite.com) logs and Reuters logs. Reuters (Reuters Ltd.) logs contain searches on a local version of AltaVista (http://www.altavista.com). For brevity purposes, in the remaining part of the paper we will refer to the two sets of logs as simply *logs*.

The Excite logs have been studied previously [5], but focused on different aspects. The logs contain transaction records of *Excite*, a major search engine. The Excite search engine does searches of documents on the exact match of the queries users entered, and it provides Boolean functions to enhance the search. The Excite logs cover all searches on *Excite* from 00:00:00 to 00:49:19 on 10th March 1997. There are 51,474 queries (or activities as referred to here) posed by 18,109 users. Each transaction record contains the following three fields:

- *Time of Day*, measured in seconds from midnight of 9 March 1997.
- *User Identification*, an anonymous user code assigned by the *Excite* server.
- *Query Terms*, exactly as entered by the given user.

The Reuters logs are transaction records of the Web searches initiated by Reuters Intranet users. The search engine used is a local version of *AltaVista* with both simple and advance search facilities. Like the Excite search engine, the Reuters search engine does searches of documents on the exact match of user queries and has Boolean functions. The time range of the Reuters logs is much wider than that of the Excite set. The logs cover from 15:27:21 on 30th March 1999 to 09:06:27 on 7th April 1999. However, the number of activities in these logs are much less than the Excite one. There are just about 9,534 activities from 1,440 unique IP addresses. Each transaction record contains the following three fields:

- *Time of Day*, measured in days, hours, minutes and seconds.
- *IP address*, a network IP address associated with the machine that the activity is initiated from.
- *CGI command*, a command line to call a CGI program for the search. It includes information about the query terms, the search method (i.e. simple search or advance search), the page number (i.e. subsequent pages for the same search), and so on.

The browsing activities are assumed to be the subsequent searches with the same queries in the Excite logs [5]. In the Reuters logs, a browsing activity is clearly marked by the page number in the CGI command. This page number can also tell us which page the user actually visited. We found 3,938 browsing activities in the Reuters logs, whose percentage among all the activities is 41.34%. This is consistent with Jansen et al's findings[5], which was 43% browsing activities, in the Excite logs.

We acknowledge that there could be browsing activities adjacent to each other that are in fact totally unrelated, which could happen in web surfing. However, the only browsing activities we found in the two logs are the requests from the user to see the returned documents in subsequent pages, which are related. This might be due to there not being any transaction records for the unrelated browsing activities. Therefore, browsing activities in this paper have a narrower meaning than the ones usually referred as *browsing* in the discussion of *web surfing*.

5 Experiments on session interval

Within the context of the experiments, we refer to **iteration** numbers of a session as the number of activities in a sequence in the session. For example, if the iteration number of a session is three, then the session has three activities. This can be a query followed by two browsing activities. We have chosen this terminology to simply emphasise the sequence in the activities within the session and their likelihood of being related to the same role.

Previous research found that most searches on the Web have just one or two iterations [5]. Generally, when the time gaps between activities are relatively small, it is reasonable to assume that the smaller the iteration number of a session, the higher chance the activities in the session are related to one role. Therefore, we would favour sessions with small iteration numbers.

A sequence of activities is grouped into a session if and only if

- the activities are from the same user ID (the Excite logs) or IP address (the Reuters logs).
- the time interval between two adjacent activities is less than or equal to the session interval in use.

By grouping the sessions with the same iteration number, we can see the distribution of various sessions. The distributions show the percentages of sessions with a particular iteration in relation to the total number of sessions. This was done because different session intervals cut the logs into different number of sessions, and so a percentage comparison was more meaningful.

Session Interval (minutes)	Percentage of 1 iteration	Percentage of 2 iteration	Percentage of 3 iteration	Percentage of 4 iteration	Percentage of 5 iteration	Percentage of 6 iteration	Sum of Percentage
1	84.07%	12.51%	2.20%	0.65%	0.26%	0.14%	99.83%
2	70.15%	17.03%	6.08%	2.93%	1.47%	0.84%	98.50%
3	65.43%	18.40%	7.12%	3.76%	1.87%	1.11%	97.69%
5	58.19%	20.02%	8.63%	5.80%	2.67%	1.67%	96.98%
10	51.66%	20.62%	9.74%	6.04%	3.57%	2.28%	93.91%
15	48.52%	20.80%	10.25%	6.46%	4.07%	2.57%	92.67%
20	47.10%	20.70%	10.38%	6.73%	4.33%	2.78%	92.02%
25	46.14%	20.65%	10.58%	6.89%	4.50%	2.86%	91.62%
30	45.31%	20.77%	10.70%	6.89%	4.63%	2.93%	91.23%
50	44.33%	20.95%	10.96%	7.25%	4.74%	2.99%	91.22%

Figure 1: The results of session interval from the Excite logs

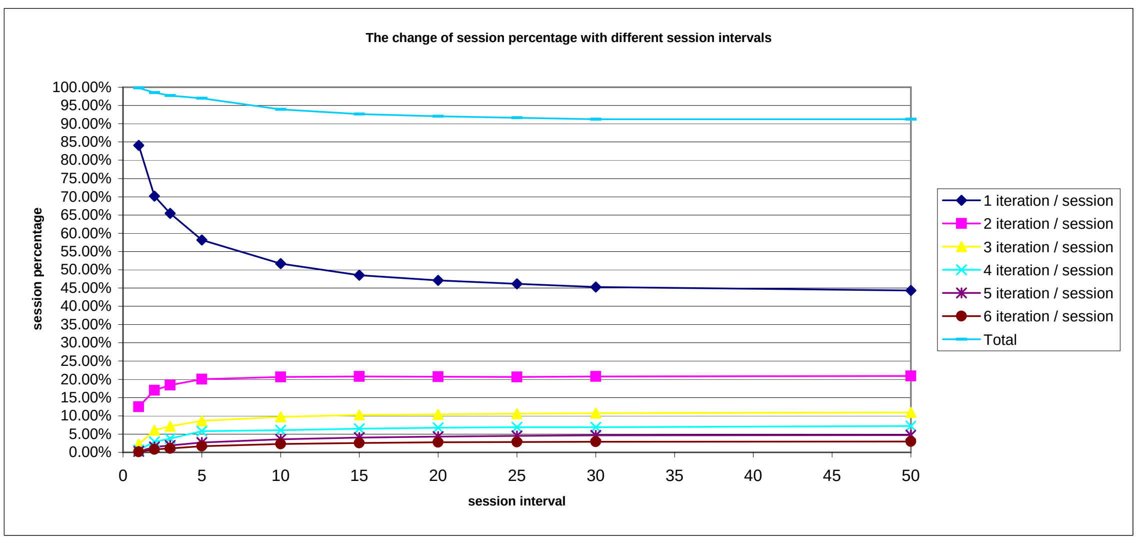

Figure 2: The results of session interval from the Excite logs

In the experiments, we first cut the logs with quite a large session interval. In the case of the Excite logs, we used a 50 minute session interval, firstly, which in fact means no session interval at all - since the time range of the Excite logs is only about 49 minutes. In the case of the Reuters logs, we used 200 minutes session interval at the beginning. Hence, we obtained initial distributions of the sessions for the two logs. Our experiment and analysis focused on monitoring only the distributions of the session with less than or equal to 6 iterations because their total covers a very large percentage of sessions ($> 91.22\%$ in the Excite logs, and $> 78.80\%$ in the Reuters logs). Then, gradually, we decreased the session interval. The aim was to cut as many large sessions in the initial distributions as possible into smaller sessions, while at the same time not affecting the small sessions (i.e. having sessions at their maximum length that contain only activities from one role - and not grouping activities from different roles). The optimal session interval is the interval that would help ensure this aim.

5.1 The results from the Excite logs

The data for the findings below, on the Excite logs, are presented in Figure 1 and illustrated in Figure 2.

- Very high percentage of sessions ($>= 91.22\%$) have less than 6 iterations irrespective of the session interval.
- More than half ($>= 65.28\%$) the sessions have 1 or 2 iterations, regardless of the session interval.
- The percentage of sessions with 1 iteration has a very small increase from the 50 minute session interval to 15 minute session interval. However, it starts to increase quickly around 15 minutes interval, and increases very dramatically when the session interval is less than 10 minutes.
- The percentages of sessions with 2 to 6 iterations are more or less stable from 50 minute session interval to the 10 minute session interval. However, they drop when the session interval is less than 10 minutes, and drop dramatically when the session interval is less than 5 minutes.

Session Interval (minutes)	Percentage of 1 iteration	Percentage of 2 iteration	Percentage of 3 iteration	Percentage of 4 iteration	Percentage of 5 iteration	Percentage of 6 iteration	Sum of Percentage
1	58.05%	12.59%	9.27%	4.56%	3.55%	2.13%	94.95%
2	51.04%	17.93%	9.92%	6.11%	4.06%	2.93%	91.99%
3	47.92%	17.82%	10.03%	7.00%	4.41%	3.22%	90.40%
5	44.18%	17.76%	10.34%	7.01%	4.77%	3.68%	87.74%
10	40.55%	17.18%	11.02%	7.57%	5.29%	3.92%	85.53%
15	38.64%	17.65%	11.82%	7.33%	5.47%	4.44%	84.35%
20	38.05%	17.33%	10.85%	7.34%	5.49%	4.59%	83.65%
25	37.51%	16.62%	11.04%	7.42%	5.74%	4.73%	83.06%
30	36.95%	16.83%	10.93%	7.38%	5.81%	4.64%	82.54%
35	36.54%	16.79%	11.04%	7.48%	5.89%	4.43%	82.17%
40	36.52%	16.62%	11.08%	7.43%	5.86%	4.37%	81.88%
60	35.67%	16.37%	10.86%	7.78%	5.93%	4.60%	81.21%
200	33.72%	15.81%	10.64%	7.64%	5.90%	4.37%	78.80%

Figure 3: The results of session interval from the Reuters logs

5.2 The results from Reuters web logs

The data for the findings below, on the Reuters logs, are presented in Figure 3 and illustrated in Figure 4.

- Very high percentage of sessions ($>= 81.21\%$) have less than 6 iterations when the session interval is less than 60 minutes.
- More than half ($>= 52.04\%$) sessions have 1 or 2 iterations when the session interval is less than 60 minutes.
- The percentage of sessions with 1 iteration has a very small increase from the 60 minute session interval to 15 minute session interval. However, it starts to increase quickly around the 15 minute interval, and it increases very dramatically when the session interval is less than 10 minutes (particularly if less than 5 minutes).

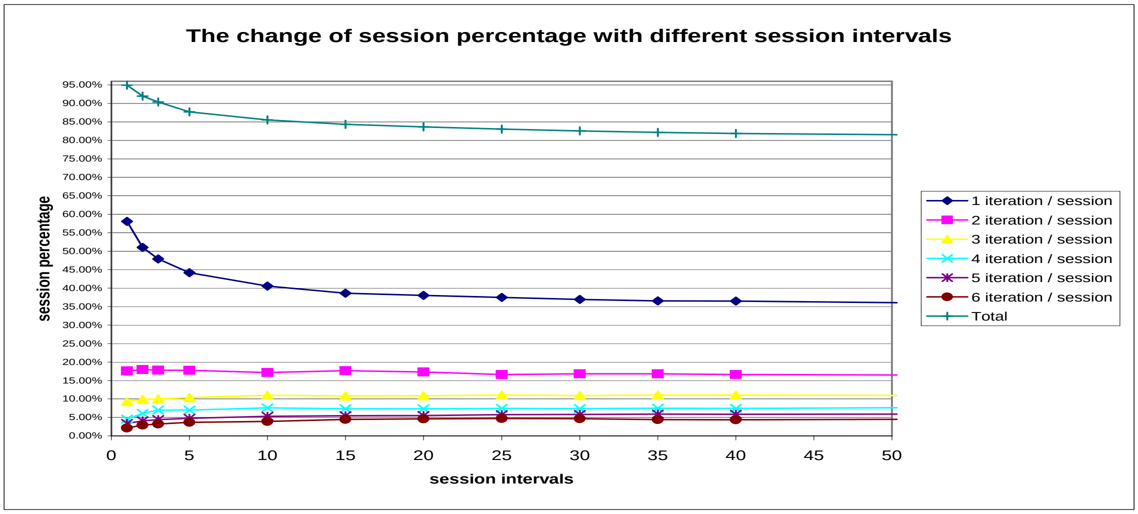

Figure 4: The results of session interval from the Reuters logs

- The percentage of sessions with 2 iterations are relatively stable from 60 minutes to 1 minute.
- The percentages of sessions with 3 to 6 iterations are almost stable from 60 minute session interval to 10 minute session interval. However, they drop when the session interval is less than 10 minutes, and drop dramatically when the session interval is less than 5 minutes.

5.3 Discussion

With two different groups of users (Internet and Intranet) in the two logs, it is interesting to see that the experiment results based on these logs are perhaps surprisingly consistent. Users of an Internet search engine (such as Excite) are believed to be much more diverse in terms of age, education, occupation, experience and aims of searching than those of an Intranet search engine (such as Reuters). Therefore, this consistency may be used as a proof for the fact that the results we obtained are valid across various users' activities.

As mentioned, the optimal session interval should not be too large nor too small. The results of the experiments show that most (short) sessions were not affected (i.e. not being cut) when the session interval is larger than 15 minutes. This probably means that the session interval is too coarse-grained to be an optimal one. When the session interval is smaller than 10 minutes, most (short) sessions are affected. In such situations, when the session interval becomes smaller, the percentage of sessions with 1 iteration increases dramatically, whereas the percentages of sessions with 3 - 6 iterations decrease dramatically. This indicates that such small session interval is too tight for most sessions. Therefore, the optimal session interval with regard to the likelihood of grouping activities from the same role together is somewhere within the range of 10 minutes to 15 minutes.

However, these results should also be confirmed with further analysis of the session cuts and the possible errors. Because a cut is totally based on the time interval, there is a possibility that several relevant search activities are divided into different sessions or several irrelevant activities are grouped together into one session. Either way is an error in the session cut. The percentage of the errors affect the applicability of the session interval. Section 6 will discuss this in details.

One immediate application of the results, and also the main motivation behind this study, is to apply them in the learning from Web users' search activities. One major obstacle in learning from Web users is the lack of relevant information about a user. For example, it is difficult to identify a user because a search engine usually does not have much information about the user unless he/she has registered, and the IP address is not a reliable resource too due to the use of proxy servers and dynamic IP allocation. However, if the logs are divided into sessions with the optimal session interval, at least the data collected within a session can be sure to be related to a particular user in a topic most of the time. So with the optimal session interval, much more data become available for the learning.

There could also be much more applications in other areas, such as studies of search interfaces, (individual and group) user modelling, and so on.

6 Errors in session cut

Using an artificial session interval to cut the activities into sessions without actually knowing more information about the activities would work most of the time if the session interval was carefully selected. However, there could be the following three types of errors:

Type 1 error: Browsing activities for the same search statement are allocated into two different sessions.

Type 2 error: Two adjacent non-browsing activities with related search statements within a short time interval are allocated into two different sessions.

Type 3 error: Unrelated searches are allocated into the same session.

Type 1 error and Type 2 error are the results of selecting a too tight session interval, whereas Type 3 error is the result of a too loose session interval. Type 2 error is vague in the sense of the time interval. It is still an open question about how short a time interval is reasonable. Considering the consequences of the three types of errors, Type 3 error is the most serious one according to our purpose because it could make the role predication invalid. Therefore, it is the type of information that is not suitable as learning data.

Having these errors does not invalidate our method of detecting sessions with time intervals. These errors indicate that people's behaviour could be irregular, so a method aiming at covering all the cases all of the time probably would not work. However, since the time interval method only aims at handling cases most of the time, this may be achieved through some careful considerations, such as using the optimal session interval.

7 Further Work: Experiments for analysing session cut errors

In order to further examine the results from selecting an optimal session interval, we need further experiments which analyse session cut errors. The aim of the experiments is to examine whether or not the percentages of errors are in an acceptable range.

Type 1 error can easily be identified in the two logs. For example,

In the Excite logs, if and only if the following relations exist between the last activity of a session and the first activity of the immediately following session, a Type 1 error is identified between the two sessions:

- the two queries have the same user ID;
- the two queries have the same query terms.

In the Reuters logs, if and only if the following relations exist between the last activity of a session and the first activity of the immediately following session, a Type 1 error is identified between the two sessions:

- the two queries have the same IP address, that is two activities from the same IP address happening in very short time are assumed to be from the same user;
- the two queries have the same query terms;
- the latter query is a browsing activity;
- the two sessions are in the same day[3].

[3] It is arguable to include this criterion. We did find that two activities satisfying the first three criteria actually happened in two different days. As it is rare, we treated it as a very special case.

Except for obvious query modifications (i.e. adding/removing words), there could be a connection between seemingly unrelated queries. Therefore, domain-specific or common sense knowledge is required for the judgement occasionally. This makes the task of recognising a Type 2 (related non-browsing activities being cut) or Type 3 error (unrelated non-browsing activities not being cut) not as straightforward as recognising a Type 1 error. All queries should be checked through manually to establish those which are related so that Type 2 and Type 3 errors can then be identified automatically when experimenting different session intervals. However, there is no guarantee that all possible errors in detecting session boundaries would be removed.

In terms of priorities, firstly the overall error (total of Type 1, 2, 3) should be as small as possible. Secondly, we view Type 3 errors to be more damaging to our cause so session intervals which minimise this type of error would be preferred.

8 Conclusion

In this paper, we argued that detecting session boundaries on the Web is important. It is problematic to assume that all the activities from the same user or IP address are in the same session, especially in logs covering several days/months. Sessions provide a common ground for various statistics relating to user activities. They also provide a possibility of regulating incremental learning when building (intra-sessional or inter-sessional) adaptive systems for the Web. Thus, it is essential that session boundaries are established on a more solid ground, such as empirical evidence.

The discussions in this paper are based on our experiments of analysing two sets of web logs: Excite logs and Reuters logs. The experiments use various session intervals, and the results show that *the optimal session interval with regard to the likelihood of belonging to one role is somewhere within the range of 10 minutes to 15 minutes.*

Our aim was to show a methodology for deriving reasonable session breaks in Web queries in order to better investigate the effectiveness of learning techniques on the Web. It is possible for learning to take place at any particular point in a session such as after/before each query, after relevance feedback, etc. so long as there is enough information for the learning purpose. Our emphasis, here, is not only on the fact that we need to have meaningful session boundaries, but also that learning techniques could better exploit these boundaries. The session boundaries provide an initial cut for learning but the sessions could also be further split if necessary. In summary, this work could provide a common terminology for discussing sessions and a mechanism for obtaining the optimal session interval.

Finally, we would like to suggest to record additional information which would facilitate future research on Web user search activities. It is beneficial to know the users' activities after the returned documents are presented to them. For example, what are the returned documents, when and which documents are selected by the users, what is the ranking information of the selected documents, and so on. If future Web search logs could provide such information, the analysis of Web users would be much easier and the results would be more accurate.

Acknowledgements

We are grateful to Reuters Ltd. and Excite Inc. for providing the Web transaction logs, Jim Jansen for his support and helpful comments about specific issues relating to TLAs, and David Harper, Josiane Mothe, Bill Teahan and three anonymous referees' invaluable comments on earlier draft of this paper.

References

[1] L. Catledge and J. Pitkow. Characterizing Browsing Strategies in the World-Wide Web. In *3rd International World-Wide Web Conference WWW95*, http://www.igd.fhg.de/archive/1995_www95/papers/, 1995.

[2] R. Cooley, B. Mobasher and J. Srivastava. Data Preparation for Mining World Wide Web Browsing Patterns. In *Knowledge and Information Systems*, 1:5-32, 1999.

[3] A. Goker and T. McCluskey. Towards an adaptive information retrieval system. In Z. W. Ras and M. Zemankova, editors, *Methodologies for Intelligent Systems. 6th International Symposium, ISMIS'91*, pages 348–357. Springer-Verlag, 1991.

[4] B. Hjorland. The concept of 'subject' in information science. *Journal of Documentation*, 48(2):182–200, 1992.

[5] Major B. Jansen, A. Spink, J. Bateman, and T. Saracevic. Real Life Information Retrieval: A Study of User Queries on the Web. *SIGIR Forum*, 32(1):5–17, 1998.

[6] S. Jones, S. Cunningham, and R. McNab. An analysis of usage of a digital library. In *Proceedings of Second European Conference on Digital Libraries*, pages 261–277, 1998.

[7] S. Robertson and C. Hancock-Beaulieu. On the evaluation of IR systems. *Information Processing and Management*, 28(4):457-476, 1992.

[8] A. Spink, J. Bateman, and Major B. Jansen. Searching heterogeneous collections on the Web: behaviour of Excite users. *Information Research*, 4(2), 1998.