International Journal of Science and Research (IJSR)
ISSN (Online): 2319-7064
Index Copernicus Value (2013): 6.14 | Impact Factor (2013): 4.438

Study on Search Behaviours Using Task Trail

P. Sreelekha[1], J. Ravikumar[2]

[1]M.Tech (CSE), Chadalawada Ramanamma Engineering College, Tirupati, India

[2]Associate Professor, Chadalawada Ramanamma Engineering College, Tirupati, India

Abstract: ***This paper analyze systematically the utilities of task level search log analysis and compare it with session and query level search log analysis in real applications. Search logs are used to determine the following: 1.Page utility estimation, 2. User search interest, 3. Website recommendations, 4. Web page rankings. Difficulty in web searches has given rise to the need for the development of personalized search engines. Personalized search engine creates user profile to capture user personal preferences. In this paper we compare Task session and Query Trails for search applications. We have applied a new concept called Task Trail which uses the clustering algorithms and performs effective search operation compared to session and query trails.***

Keywords: Session trail, Query trail, Task trail, Log analysis

1. Introduction

The searching activities of users are recorded by web search log. Web search log can be used in various applications like
1) Prediction of user search interest
2) Website recommendation
3) Web page re-ranking methods
4) Query Suggestion.

In this paper we compare the performances of session, query, and task trails in applications including:
a) Determination of user satisfaction where dwell time(amount of time between click and action) and success course of markov models are mined to measure user satisfaction
b) Prediction of user search interest where ODP category information is used to measure topic similarity.
c) Measurement of ranking functions where the difference of two ranking functions in session, query, and task level are measured.

Dwell Time
It is the amount of time between the click and next action. This time is a good indicator for user satisfaction. The more the dwell time the more is the success of the search. The behaviour of set of users in the web log may be either search behaviour or browse behaviour. Search behaviour is a single query submitted to the search engine.Browse behaviour may be one of the following:
1) Starting to surf from the home page.
2) Typing a URL address
3) Pasting a URL address from another page to the address bar of the existing webpage.
4) User clicks a bookmark or a back or forward button in a browser.

Log Segmentation can be done in any one of the following trails:

Query Trail
It represents a sequence of user behaviours of one of the user starting from a query followed by sequence of browsing behaviours that are triggered by this query.

Session Trail
It represents a sequence of user behaviours of one of the user where user behaviours are consecutive and any two consecutive occurred within the time threshold.

Disadvantages of existing system
In case of query trails the semantic association between adjacent query trails are lost. In case of session trails it strictly follows the chronological order of user behaviours in search logs. Time threshold settings for session segmentation are not able to satisfy our predefined goals. Sessions contain multiple atomic information need which are irrelevant semantically. ODP category information is required to predict user search interest. Co-occurrence based query suggestion methods based on tsk trail are compared with same methods based on session trails and click through bipartite graph.

2. Literature Survey

In 2010, B. Xiong, Z. Liao conducted study on real search logs and developed 4 principles for ranking based context-aware.

In 2010 R. Jones, A. Hasson addressed the problem of predicting user search logs. They explained that user behaviour can explain the success of user search goals.

In 2010 N. Craswell compared the performances as measured by judgement-based information and usage-based information.

In 2010 R. White, A. Single compared different methods for finding the best trail.

In 2011 Y. Sun, A. Hassan analysed user behaviour to predict whether the user ended up being satisfied with search or not.

In 2011, S. Silver Striz, S. Orlondoy, C. Lucchesez proposed a clustered based algorithm and introduced Query time-stamps.

In 2012, O. Chapelle, F. Radlinski analysed interleaving data.

In 2012, D. Zhoo used the user pattern to build suggestion models.

In 2013, J. Allan, H. Field analysed Dwell times, page visits for recommendations including user preferences.

In 2014, Zhen Liao introduced Task Trail to understand user search behaviour, predicting search interests and suggesting related queries.

3. Proposed System

To overcome the draw backs of session and query trails we have introduced Tsk trails.

Task Trail:
It represents a sequence of user behaviours of one user within one session where all user behaviours collectively define an atomic user information need.

Task segmentation contains two steps:
1. Logs are segmented into sessions based on threshold.
2. Segment session into tasks based on semantic relationships between queries.

Advantages
1. Task trail performs better than session and query trail in determination of user satisfaction.
2. Page utilities can be increased by Task trail.
3. In measuring ranking function, Task trails are more sensitive than session trails and comparable to query trails.
4. Topic similarity can be well preserved by task trail as because it provides atomic user information needs.

Clustering Framework

Entire query trail can be represented by starting search query of each query trail. We can group two queries into the same task if a) Two queries are identical b) One is a part of the other. c) Two partially agree to each other. d) One is the type of the other. The basic ideas of clustering framework are
1) Logs are segmented into sessions by using some threshold.
2) Similarity between two queries are computed.
3) Similar queries are clustered into the same task.

We construct an undirected graph for queries. Vertices indicate queries and an edge indicates similarity scores between queries. The edges having the score less than threshold are suspicious edges and they are to be removed. After removing the suspicious edges any connected componentof the remaining graph is identified as a task. This approach is called query clustering using weighted connected component of a graph (QC-WCC). However this method has disadvantage that it has high time complexity for constructing the graph [O (K.N^2)] were N is the numbering of queries and K is the dimension of features. Computing the pair wise similarity for all consecutive query pairs is the better approximation.

Head tail component query clustering (QC-HTC) uses the heuristics that queries are submitted sequentially by users. It violates the task interleading observations found by us in search logs. So an algorithm called query clustering using bounded spread method (QC-BSP) is used.

Algorithm 1: Spread Query Task Clustering (QC-SP).

Input: Query set Q, cut-off threshold z;
Output: A set of tasks Θ;
Initialization: Θ = ∅; L=∅;
1: **for** *len*= 1 :|*Q*| − 1 **do**
2:**for** *i* = 1 :|*Q*| − *len***do**
3: // if two queries are not in the same task
4: **if** $L[Qi] \neq L[Qi+len]$ **then**
5: // compute similarity is $O(k)$
6: $s \leftarrow \mathrm{sim}(L[Qi];L[Qi+len])$;
7: **if** $s \geq b$ **then**
8: merge Θ (*Qi*) and
Θ (*Qi*+*len*);
9: modify *L*;
10: // break if there is only one task
11:**if** | Θ |= 1 break;
12: return Θ;

By the above algorithm we observed that consecutive query pairs are more likely belonging to same tasks compared to non-consecutive ones. In this we will calculate the similarities for consecutive query pairs. Consider an example consisting of series of queries $\{q_1, \rightarrow q_2, \rightarrow q_3, \rightarrow q_4\}$. This algorithm will calculate for pairs $\{q_1\ q_2, q_2\ q_3, q_3\ q_4\}$This will take a time complexity of O(K.N) if the sequences $\{q_1, q_2, q_3, q_4\}$ are grouped into $\{q_1\}$ and $\{q_2, q_3, q_4\}$ the standard approach will compute all six query pairs but QC-BSP needs to calculate for five query pairs only. $\{q_2\ q_3\}$ is skipped since the query pairs are similar.

Effectiveness of different clustering algorithms

Method	Time (sec.)	RI	JI	HI
QC-WCC	3,093	1.000	1.000	0.528
QC-HTC	1,001	0.949	0.899	0.519
QC-SP	1,902	1.000	1.000	0.528
QC-BSP-3	242	0.939	0.855	0.533
QC-BSP-5	418	0.966	0.919	0.531
QC-BSP-10	807	0.988	0.972	0.529

4. Results

Consider two data sets D_0 and D_1. D_0 consistsof user browsing logs from a widely used browser plug-in. D_1 consists of web search logs from bing. It contains information about user anonymized unique identifier (machine ID), unique browser identifier, user clicked URL and queries related to clicks, a referral URL and time stamp.

Multitask and interleaved task behaviour can be illustrated by using different threshold to extract session ranging from one minute, two minutes, five minutes, ten minutes, twenty minutes, thirty minutes, sixty minutes, to a day. The following table illustrates 1) multitask behaviour always exists in search processes. 2) Interleaved tasking behaviour may not occur if threshold is less than five minutes.

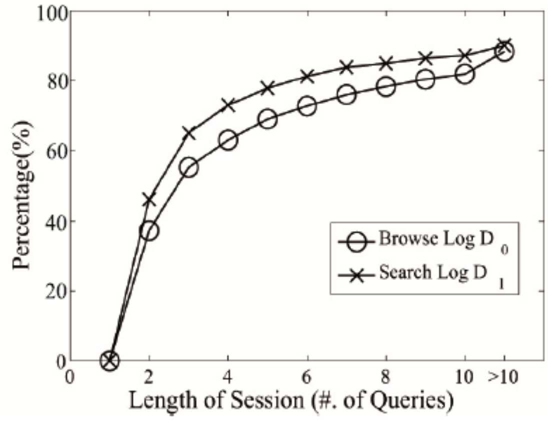
Statistics on browse and search logs

Statistics	D_0	D_1
Avg. # of Queries in Sessions	5.81	2.5
Avg. # of Queries in Tasks	2.06	1.6
Avg. # of Tasks in Sessions	2.82	1.5
% of Multi-Task Sessions	42.6	28.8
% of Interleaved Task Sessions	12.6	4.41
% of Single-Query Tasks	48.7	71.8
% of Multi-Query Tasks	51.2	28.1

Sessions in D_0 are longer than that of D_1 because users often browse before and after searching.

The following graph helps in observations of length of session with respect to percentage of session.

The graph consists of search logs and browse logs. Browsing is done by the user before and after searching.



Increase in length of session will decrease the percentage of session.

Analysis of User satisfaction

The following feedback signals will help us inanalysing the user satisfaction

Clicks:
Users' success can be determined by the time between the click in user search goal and the action

Dwell Time:
Long Dwell time is a predicator of success. It is the amount of time between click and next action.

Success Scores:
Markov models are used to determine the success scores. Two Markov models are used to determine the likelihood of user satisfaction and dissatisfaction.

Ranking Function

If the Ranking function A is better than another Ranking function B, it implies that user satisfaction is more at task level.

Sensitivity is calculated by P-value (T-test). It includes the following steps

1) Collect the average user satisfaction rates in both rankers.
2) Conduct sample T-test.
3) Get a probability that user satisfaction rate is same for Ranker A and Ranker B

Table: Effectiveness of different Ranking Functions

Measure	Session	Task	Query
Click Rate	0.0796	0.0164	0.0563
30S Click Rate	0.0276	0.0186	0.0192
MM Success Rate	0.0031	0.0016	0.0014

From the above table we observe the following:

a) Among different implicit signals MM Success rate has smallest P-Value which indicates that MM Success rate is more sensitive than other two measures.
b) Using Click Rate as implicit measure is not good. Because Ranking functions are same at Task,Query, Session levels.
c) Using 30S Click Rate and MM Success rate, Task level measurements are more sensitive than Session and Query level.

5. Conclusion

Our study developed to use Task Trail for better and efficient understanding of search behaviours. Our Analysis and comparisons approved the effectiveness of Task Trail in several search applications including estimation of user satisfaction, prediction of search interest of users. The algorithm that we used will reduce the time complexcity to merge similar queries into same task. Also Topic similarity between the query pairs is well preserved.We analysed the effectiveness of all the Session, Query, and Task levels trails and found that users are more interested to find useful information from Task Trail.

References

[1] Chapelle O, Yisong,"Large scale validation and analysis of interleaved search evaluation" ACM Trans. Inf. Syst. Vol 1, p.6.
[2] Hassan, Song., "A task level user satisfaction metric and its application on improving relevance estimation" ser. CIKM '11.
[3] Vapnik v., "Staticial learning theory"
[4] Liao, Z.Song, Ëvaluating the effectiveness of search trails ser. WWW 12, 2012, pp. 489-497.
[5] Jones, K.L, "Beyond the session time out: automatic hierarchical segmentation of search topics in query logs, CIKM '08 ACM, 2000.
[6] F.Radlinski and Craswell, "comparing sensitivity of information retrieval metrics", in proc 33rd intconf res. Develop.
[7] White Bilenko,Crcerzan study the use of popular destination to chance web search interactions. Proc. SIGR, 159-196.
[8] White, S.M(2007) investigated behavioural variability in web search proc. WWW,21-30.
[9] D-Z.hov and L-W.He, "query suggestion by constructing term transaction graphs", in proc 5th ACM int. conf. web search DM,2012, pp. 353-362.
[10] H-Wang, W.chv, M-W chang, "learning to extract cross-section search tasks" in proc 22nd int. conf. WorldWideWeb,2013,pp. 1353-1364.
[11] Catledge and pitkow ,"characterizing browsing strategies"in World Wide Web", pp. 1065-1075.
[12] Jain, Velipasaaoglu E, "Synthasising high utility suggestion for rare web search queries, "SER-SIGIR", 2011, pp. 805-814.