

Contextual and Content Based User Identified Web Re-Visitation

Firoj Alam¹, Md Ateeq Ur Rahman²

¹Research Scholar, Dept. of Computer Science & Engineering, SCET, Hyderabad

²Professor and Head, Dept. of Computer Science & Engineering, SCET, Hyderabad

Abstract- This paper leverages human's natural recall method of victimisation episodic and long-term memory cues to facilitate recall, and presents a private internet revisit technique referred to as WebPagePrev through context and content keywords. Getting back to antecedently viewed websites may be a common nevertheless uneasy task for users thanks to the big volume of in person accessed data on the online. once a user interacts with WebPagePrev during net revisit section, we will either manually enter some context keywords, or obtain instructed values from discourse hierarchies by clicking the leftside buttons of your time, location, and activity bars . Underlying techniques for context and content memories' acquisition, storage, decay, and utilization for page re-finding square measure mentioned. A connection feedback mechanism is additionally concerned to tailor to individual's memory strength and revisit habits. Our study shows that, Compared with the present internet revisit tool souvenir, History List looking methodology, and computer program methodology, the projected WebPagePrev delivers the simplest re-finding quality to find rate, average F1-measure and average rank error . Our dynamic management of context and content recollections as well as decay and reinforcement strategy will mimic users' retrieval and recall mechanism. With connection feedback, the finding rate of WebPagePrev will increase, average F1-measure will increase, and average rank error decreases compared to stable memory management strategy. Among time, location, and activity context factors in WebPagePrev, activity is that the best recall cue, and context+content primarily based re-findingdelivers the simplest performance, primarily compared to context based re-finding and content based re-finding.. Relevance feedback is associate interactive approach that has been shown to figure significantly well in classical data retrieval and additional recently in net search domain. Every discourse hierarchy is dynamically maintained by analyzing the user's clicking behaviors and the applied math frequencies of captured context instances. Frequently accessed context things square measure prime listed within the corresponding discourse hierarchy. When a user accesses a web page, that is of potential to be revisited later by the user i.e., page time interval is over a threshold, the context acquisition and management module captures the current access context i.e., time, location, activities inferred from the presently running pc programs into a probabilistic context tree. Meanwhile, the content extraction and management module performs the unigram based extraction from the displayed page segments and obtains a

listing of probabilistic content terms. the possibilities of non-heritable context instances and extracted content terms replicate however doubtless the user can sit down with them as memory cues to induce back to the antecedently centered page. As well we deal with context and content ambiguity considering confusion and error throughout memory construction.

Index Terms- Web revisit, LTM ,access context, page content, relevance feedback, CTMSP-Mine.

I. INTRODUCTION

To Nowadays, the online is enjoying a major role in delivering information to users' fingertips. an internet page is localized by a hard and fast address, and displays the page content as time-varying photograph.

To deal that considering access context, users may retreat to the previous viewed pages through some content keywords. Instead of extracting content terms from the total net page, we have a tendency to solely contemplate the page segments shown on the screen. There area unit several term coefficient schemes in the information retrieval field. the foremost generic one is to calculate term frequency-inverse document frequency (tf-idf) [36]. For customized net revisit, merely counting the prevalence of a term within the given pagesegment isn't enough. Also, user's web content browsing behaviors (e.g. visitation time length and light or not), yet as page's subject headings, area unit counted as user's impression and potential interest indicators for later recall. in an exceedingly similar manner as access context, we bind a bearing score to every extracted content term d, showing however doubtless the user can seek advice from it for recall based on the four normalized options Among the common internet behaviors, web revisit is to re-find the antecedently viewed web pages, not solely the page address, however conjointly the page snapshot at that access timestamp . A 6-week user study with twenty three participants showed nearly fifty eight of internet access belonged to internet revisit. Another 1-year user study involving 114 participants unconcealed around 40% of queries were re-finding requests [3]. in keeping with [4], on average, each second page loaded was already visited before by identical user, and also the magnitude relation of revisited pages among all visits ranges between two hundredth and seventy two. Psychological studies show that humans have faith in each LTM and LTM to recall data or events from the past. Human's episodic memory receives and stores temporally dated episodes or

events, in conjunction with their spatial-temporal relations, while human's LTM, on the opposite hand, is a structured record of facts, meanings, ideas and skills that one has nonheritable from the external world. Semantic information comes from accumulated episodic memory. LTM is thought of as a "map" that ties along things in LTM. The two reminiscences frame the class of human user's declarative memory, and work along in user's data recollecting activities [5]. Thus, once a user's web revisit behavior happens, s/he tends to utilize episodic memory, interwoven with LTM, to recall the antecedently targeted pages. Here, semantic memory accommodates content data of antecedently focused pages, and LTM keeps these pages' access context (e.g., time, location, concurrent activities, etc.) [6], [7]. Inspired by the psychological findings, this paper explores how to leverage our natural recall method of victimisation episodic and LTM cues to facilitate personal web revisit. Considering the variations of users in memorizing previous access context and page content cues, a connexion feedback mechanism is concerned to enhance personal internet revisit performance.

II. RELATED WORKS

Existing System

The services which are provided to the wireless mobile devices (such as PDAs, Cellular Phones, and Laptops) from anywhere, at any time using ISAP (Information Service and Application Provider) are enhanced by mining and prediction of mobile user behaviors. But such discovery may not be precise enough for predictions since the differentiated mobile behaviors among users and temporal periods are not considered simultaneously in the previous works. User relations and temporal property are used simultaneously in this work. Prediction strategy is used to predict the subsequent mobile behavior.

Here CTMSP-Mine (Cluster-based Temporal Mobile Sequential Pattern - Mine) algorithm is used to mine CTMSPs. In CTMSP-Mine requires user clusters, which are constructed by Cluster-Object-based Smart Cluster Affinity Search Technique (CO-Smart-CAST) and similarities between users are evaluated by Location-Based Service Alignment (LBS-Alignment) to construct the user groups. The temporal property is used by time segmenting the logs using time intervals. The specific time intervals to segment the huge data logs are found using Genetic Algorithm based method called GetNTSP (Get Number of Time Segmenting Points). The user cluster information resulting from CO-Smart-CAST and the time segmentation table are provided as input to CTMSP-Mine technique, which creates CTMSPs. The prediction strategy uses the patterns to predict the mobile user behavior in the near future.

III. PROPOSED SYSTEM

From user's computer programs. Instead of extracting content items from the full web page as done in , we extract them from page segments displayed on the screen in the user's view, and assign a probabilistic value to each

extracted term based on user's page browsing behaviors (i.e., dwell time and highlighting), as well as page's subject headings and term frequency-inverse document frequency (tf-idf), reflecting user's impression and likeliness of using the keyword as recall content cues. Other closely related work such as enabled users to search for contextually related activities (e.g., time, location, concurrent activities, meetings, music playing, interrupting phone call, or even other files or web sites that were open at the same time), and find a target piece of information (often not semantically related) when that context was on. This body of research emphasizes episodic context cues in page recall. How to grasp possibly impressive semantic content cues from user's page access behaviors, and utilize them to facilitate recall are not discussed. To tailor to individual's web revisit characteristics, as well as human user's context and content memory degradation nature, this study presents methods to dynamically tune influential parameters in building and maintaining probabilistic context and content memories for recall.

Synthetic Data Generation We firstly build two extra components: 1) data simulator, to simulate the generation of personal photo collections; and 2) user simulator, to simulate the user's memory over the generated data and revisit actions, acts as a "real user". Synthetic data generation lies in the following two aspects. Generation of Context and Content Memory. The data simulator crawls users' with contextual metadata (e.g., time, location) to form a dataset using Flickr API⁷ from social network. Considering personal photo collection, data simulator mainly selects users who share more than 500 photos on Flickr. Then the data simulator generates photonum (1 k, . . . , 100 k) photos, which correspond to context lattices and content term lists to mimic memory snapshot. Generation of Revisit Requests. Every period (seven days), the user simulator formulates 10 revisit requests against above generated context and content memory. Each revisit request contains keyword num (2, . . . , 10) keywords, which are randomly selected from the corresponding context lattices and content term lists. PhotoPrev processes the revisit requests from the user simulator periodically, and then relevant parameters are updated based on user feedback adaptation. The obtained probabilistic context trees can evolve dynamically in life cycles to replicate the gradual degradation of human's episodic learning in addition because the context keywords that users can use for recall. That is, for each node within the probabilistic context tree, its association score can increasingly delay with time.

IV. SYSTEM ARCHITECTURE

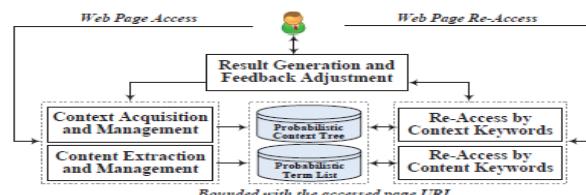


Fig.2: The personal web revisit framework

In the literature, a number of techniques and tools like bookmarks, history tools, search engines, metadata annotation and exploitation, and contextual recall systems have been developed to support personal web revisit. The most closely related work of this study is Memento system [8], which unifies context and content to aid web revisit. It defined the context of a web page as other pages in the browsing session that immediately precede or follow the current page, and then extracted topic-phrases from these browsed pages based on the Wikipedia topic list. In comparison, the context information considered in this work includes access time, location and concurrent activities automatically inferred from user's computer programs. Instead of extracting content items from the full web page as done in [8], we extract them from page segments displayed on the screen in the user's view, and assign a probabilistic value to each extracted term based on user's page browsing behaviors (i.e., dwell time and highlighting), as well as page's subject headings and term frequency-inverse document frequency (tf-idf), reflecting user's impression and likeliness of using the keyword as recall content cues. Other closely related work such as [9], [10] enabled users to search for contextually related activities (e.g., time, location, concurrent activities, meetings, music playing, interrupting phone call, or even other files or web sites that were open at the same time), and find a target piece of information (often not semantically related) when that context was on. This body of research emphasizes episodic context cues in page recall. How to grasp possibly impressive semantic content cues from user's page access behaviors, and utilize them to facilitate recall are not discussed. To tailor to individual's web revisit characteristics, as well as human user's context and content memory degradation nature, this study presents methods to dynamically tune influential parameters in building and maintaining probabilistic context and content memories for recall.

A. Module Description:

Modules:

1. Web user Module
2. Admin Module
3. Chart module
4. Image upload module
5. Content Extraction and Management Module:

1. web user Module :

The web user (User) is allotted a world user identity Uid . user ought to register page before gonna to login the page.while login user ought to provide location as a result of supported location we are able to search the product,after login user will search any product,they can read the similar product , similar word recommendation facilitate for user.

2. admin Module:

admin module, admin is super user,admin will read all the details,here admin will read the chart supported user product revisit,user will add the look details and merchandise details supported location,semantic memory cues to

facilitate recall, and presents a private internet revisit technique referred to as WebPagePrev through context and content keywords.

3. chart module:

chart module,chart module supported range times product return by some one,so user will simply conclude that product move terribly quickly its all supported user looking,User relations and property area unit used at the same time during this work. Prediction strategy is employed to predict the next mobile behavior

4. image upload module:

Upload image module, admin will transfer any image , once uploading image user will give correct authentication details they'll read file This paper leverages human's natural recall method of victimization episodic and LTM cues to facilitate recall, and presents a private net revisit technique referred to as WebPagePrev through context and content keywords.

5. Content Extraction and Management Module:

Apart from access context, users may come back to the previous viewed pages through some content keywords. Instead of extracting content terms from the total internet page, we tend to solely contemplate the page segments shown on the screen. There ar several term weight schemes in the information retrieval field. the foremost generic one is to calculate term frequency-inverse document frequency (tf-idf) [36]. For personalised internet revisit, merely counting the incidence of a term within the bestowed page segment isn't enough. Also, user's website browsing behaviors (e.g. visitation time length and lightness or not), additionally as page's subject headings, ar counted as user's impression and potential interest indicators for later recall. during a similar manner as access context, webind a sway score to every extracted content term d, showing however seemingly the user can talk to it for recall based on the four normalized options.

V. OBSERVATION

When a user will re-finding, s/he sometimes has bound purposes in mind, like making ready a project proposal, writing codes, etc. WebPagePrev strives to support users to re-find what they accessed through previous access time, location, synchronal activities, and content keywords. Beyond that, additional user-centric context factors (e.g., access purpose, expertise, background, interest, etc.), in addition as social context factors (e.g., external events, encompassing individuals, etc.), can be inferred from user's profile, agenda, and external service suppliers, and delimited with the accessed pages. during this manner, not only the user him/herself may benefit from such wealthy contextual cues throughout re-finding method, however conjointly different users with similar access purpose and background might share the additional directed page access. this can be in line with the spirit of social search that

advocates two paradigms (namely, library paradigm and village paradigm) in info retrieval. consistent with, in a library, individuals use keywords to go looking documents, and the trust relies on authority, whereas during a village, people use language to raise queries, answers are generated in period by anyone with the experience in the community, and trust relies on intimacy.

In social search, plenty of information regarding the individuals is employed, bringing in privacy protection problems. Life-cycle management of people's info with degradation policies from high to low exactness, as finished the context memory mechanism during this study, can be exploited. We leave this issue to our additional study.

VI. CONCLUSION

In this work, we tend to planned a technique to mechanically construct AN adaptational and evolutive context and content memory supported users' personal pic collections, supporting users' pic revisitation by keywordbased queries on sensible phones. The planned methodology is evaluated by AN experiment on an outsized artificial dataset and a 12-week user study. Our experimental results show that it will adapt to the user's return habit and supply a straightforward nevertheless effective answer exploitation human memory cues. As future work, we'd wish to deal with context and content ambiguity considering confusion and error throughout memory construction In this work, we have a tendency to planned a technique to mechanically construct AN adjustive and evolutive context and content memory supported users' personal pic collections, supporting users' pic revisitation by keyword based queries on sensible phones. The planned methodology is evaluated by AN experiment on an outsized artificial dataset and a 12-week user study. Our experimental results show that it will adapt to the user's come back habit and provide an easy nevertheless effective answer mistreatment human memory cues. As well we deal with context and content ambiguity considering confusion and error throughout memory construction.

VII. BIBLIOGRAPHY

- [1]. L. Tauscher and S. Greenberg. Revisitation patterns in world wide web navigation. In CHI, pages 399–406, 1997.
- [2]. A. Cockburn, S. Greenberg, S. Jones, B. Mckenzie, and M. Moyle. Improving web page revisitation: analysis, design and evaluation. IT & Society, 1(3):159–183, 2003.
- [3]. 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(1):97–137, 1997.
- [4]. J. Teevan, E. Adar, R. Jones, and M. Potts. Information retrieval: repeat queries in yahoo's logs. In SIGIR, pages 151–158, 2007.
- [5]. M. Mayer. Web history tools and revisitation support: a survey of existing approaches and directions. Foundations and Trends in HCI, 2(3):173–278, 2009.
- [6]. T. Deng, L. Zhao, H. Wang, Q. Liu, and L. Feng. Refinder: a context-based information re-finding system. IEEE TKDE, 25(9):2119–2132, 2013.
- [7]. T. Deng, L. Zhao, and L. Feng. Enhancing web revisitation by contextual keywords. In ICWE, pages 323–337, 2013.
- [8]. H. Takano and T. Winograd. Dynamic bookmarks for the WWW. In HYPERTEXT, pages 297–298, 1998.
- [9]. S. Kaasten and S. Greenberg. Integrating back, history and bookmarks in web browsers. In HCI, pages 379–380, 2001.
- [10]. J. A. Gamez, J. L. Mateo, and J. M. Puerta. Improving revisitation browsers capability by using a dynamic bookmarks personal toolbar. In WISE, pages 643–652, 2007.
- [11]. R. Kawase, G. Papadakis, E. Herder, and W. Nejdl. Beyond the usual suspects: context-aware revisitation support. In HT, pages 27–36, 2011.
- [12]. D. Morris, M. R. Morris, and G. Venolia. Searchbar: a searchcentric web history for task resumption and information refinding. In CHI, pages 1207–1216, 2008.