IOSR Journal of Computer Engineering (IOSR-JCE)
e-ISSN: 2278-0661,p-ISSN: 2278-8727, Volume 20, Issue 3, Ver. III (May. - June. 2018), PP 1-6
www.iosrjournals.org

Web User Revisitation Relevance Feedback Using Keyword For Web Mining

[1]Ruchi Patil[2]Prof. NitinRaut
Gurunanak Institute of Engineering & Technology, Dahegoan, Nagpur, India
Gurunanak Institute of Engineering & Technology, Dahegoan, Nagpur, India
Corresponding Author: Ruchi Patil

Abstract: *Nowadays, the web is playing a significant role in delivering information to users' fingertips. A web page can be localized by a fixed url, and displays the page content as time-varying snapshot. Among the common web behaviors, web revisitation is to re-find the previously viewed web pages, not only the page url, but also the page snapshot at that access timestamp. A 6-week user study with 23 participants showed nearly 58% of web access belonged to web revisitation. Another 1-year user study involving 114 participants revealed around 40% of queries were re-finding requests. According to, on average, every second page loaded was already visited before by the same user, and the ratio of revisited pages among all visits ranges between 20% and 72%.. In this paper*

Date of Submission: 02-06-2018 Date of acceptance: 18-06-2018

I. Introduction

As far back as they were made, web indexes are utilized tofind URLs on the web, as per given keywords. They havebeen a helpful mechanism to researchers everywhere throughout the world, as the exponential development of the Internet ended up plainly overpowering. Be that as it may certain issues emerge. In particular, as most web indexes don't consider client interest, the URLs came back from them are the same when diverse clients input a similar question words. It brings about the greater part of data returned irrelevant to client interest [1].The Web Mining research are has a place with Mining research are has a place with Several Groups Such as the Database, Artificial Intelligence and Information Retrieval. We Categorize Web Mining into three regions: Web Structure Mining, Web Content Mining, and Web Usage Mining. Web content mining centers around the retrieval and Discovery of the helpful information substance or information or records from the Web [2]. Web structure stresses to the disclosure of how to show the hidden connection structures of the Web. As the Internet information increments drastically, the general inquiry mode, not considering of clientintrigue, is increasingly hard to address the issues of individuals in information recovery. In the meantime, the investigation oncustomized internet searcher stands out enough to benoticed. We propose some new customized strategiesapplying to the web index.Web is an expansive Dynamic, Different and UnstructuredData Repository field. Web an expansive, hazardous various, dynamic and generally unstructured informationRepository which suppliestremendous measure of data [3].Existing web search toolsfor example, Google, Yahoo and Amazon frequently restorea not insignificant rundown of query items, positioned bytheir pertinent relationships to the predetermined question.Web clients need to experience the rundown and inspect thetitles, labels and (short) bits successively to recognize theirrequired outcomes.This is a tedious errand since varioussubtopics of the given inquiry are combined.This paperexhibits the expression and updated strategies for clientinterests in light of client search histories, and gives a filteralgorithm in light of this model [3]. Moreover, engineeringof a new personalized search engine framework is outlined,and its principle thoughts are described. The architecture cangivepersonalized search comes about for various clients, andmake strides search comes about viably, and have greatflexibility.

II. Related Work

Deng et.al. built a unified index of information that a person has seen on the computer, including emails, web pages, documents, media files, calendar appointments, etc., and allowed the use of such contextual cues as time, author, thumbnails and previews to search for information. Deng et al. allowed users to re-find web pages and local files through previous access context, including time, location, and concurrent activities. It clustered and organized context instances in a context memory, and dynamically degraded the context instances to mimic user's memory decay feature.

A query-by-context model for information recall was presented upon the context memory. YouPivot leveraged human user's natural method of recall by allowing a user to search through their digital history (e.g.,

files, URLs, physical location, meetings and events) for the context they do remember. On the other hand, there exist systems with moderate degree of content similarity in their primary storage such as email servers, virtualized servers, and devices .The user can then Pivot, or see everything that was going on while that context was active. Further, YouPivot displayed a visualization of the user's activity, providing another method for finding context. Memento providare included. We also show that fixed length chunks work well, achieving nearly the same compression rate as variable-length chunks. Finally, we show that simply identifying zero-filled blocks, even in ready-to use virtual machine disk images available online, can provide significant savings in storage.

J. Li and W.Wanget. al. Haystack stored arbitrary objects of interest to a user, and recorded arbitrary (predefined or user-defined) properties of and relationships between the stored information. It coined a uniform resource identifier (URI) to name anything of interest, including a document, a person, a task, a command/menu operation, or an idea.

Once named, the object can be annotated, related to other objects, viewed, and retrieved through arbitrary properties, which served as useful query arguments, as facets for metadata-based browsing, or as relational links to support the associative web browsing.

Bearing the similarity to Haystack, a SQLbasedMyLifeBitsplatform was built for recording, storing, and accessing a personal lifetime archive. It stored content and metadata for a variety of item types, including contacts, documents, email, events, photos, music and video, which were linked together implicitly using "time", or explicitly linked with typed links such as a "person in photo" link between a contact and a photo,or a "comment" link between a voice comment and a document. With linking, the traditional folder (directory) tree was replaced by a more general "collections" function using a directed acyclic graph (DAG).

III. Proposed System

The proposed system categories into following modules

1. **Web User Module**
2. **Admin Module**
3. **Chart Module**
4. **Image Upload Module**

Web User Module:

The web user (User) is assigned a global user identity Uid .User should register page before going toto login the page. While login user should give location because based on location we can search the productsafter login user can search any product,they can view the similar product, similar word recommendation help for user.

Admin module:

Admin module, admin is super user,admin can view all the details,here admin can view the chart based on user product revisitation,user can add the shop details and product details based on location,semantic memory cues to facilitate recall, and presents a personal web revisitation technique called WebPagePrev through context and content keywords.

Chart module:

Chart module, chart module based on number times product revisit by someone,so user can easily find out which product move very quickly it's all based on user searching, User relations and temporal property are used simultaneously in this work. Prediction strategy is used to predict the subsequent mobile behavior.

Image upload module:

Upload image module, admin can upload any image , after uploading image user can provide correct authentication details they can view file This paper leverages human's natural recall process of using episodic and semantic memory cues to facilitate recall, and presents a personal web revisitation technique called WebPagePrev through context and content keywords.

Flowchart

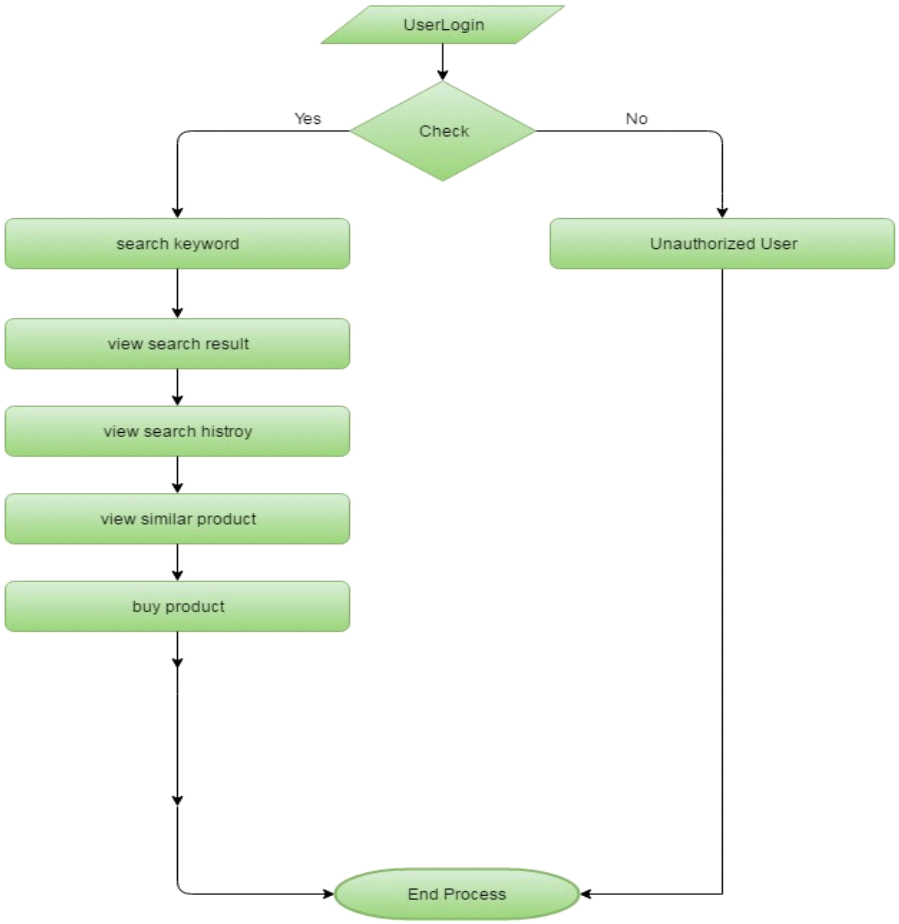

Figure 1: Flowchart of proposed system

IV. Results Analysis

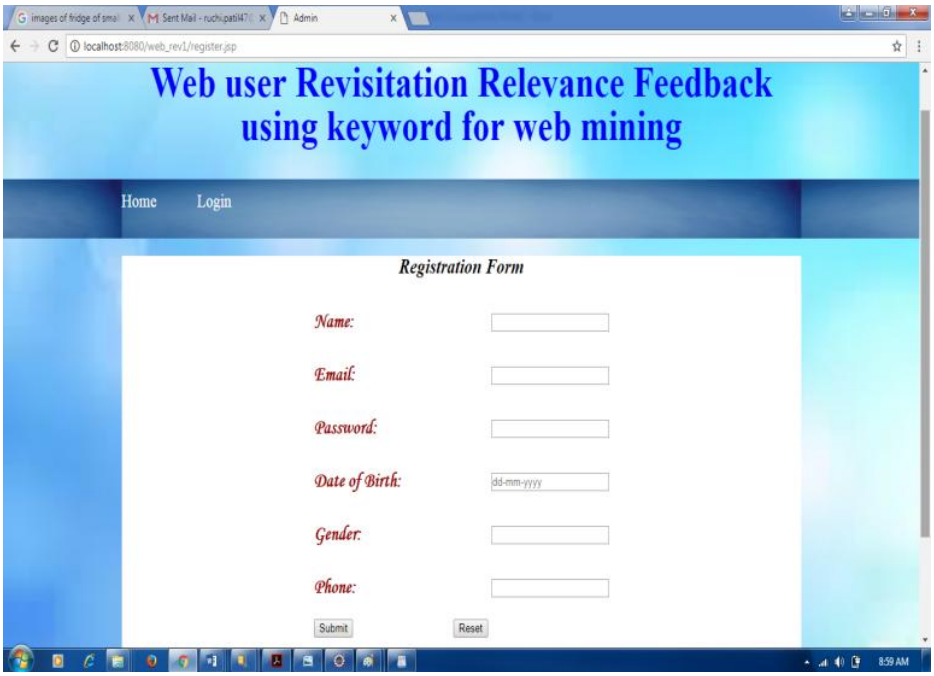

Figure 2: Registration form

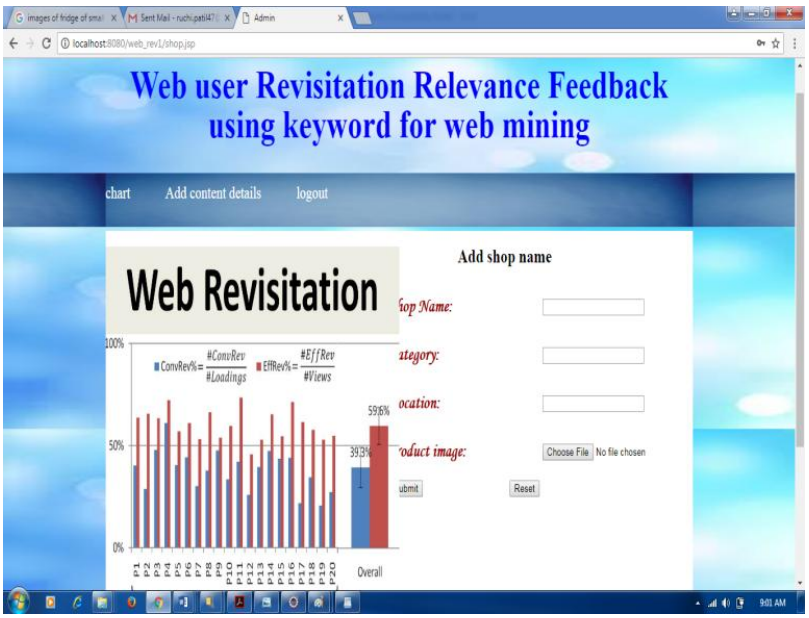

Figure 3: Admin page for adding shops

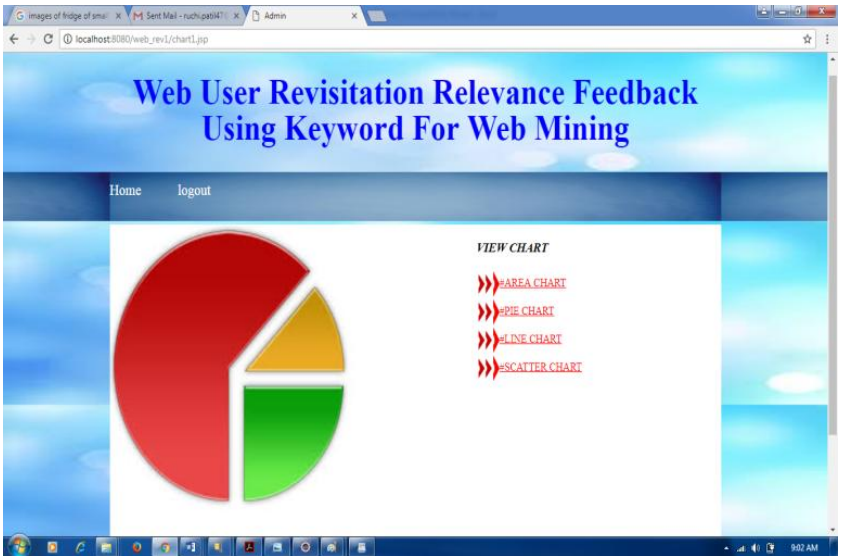

Figure 4: Chart Page

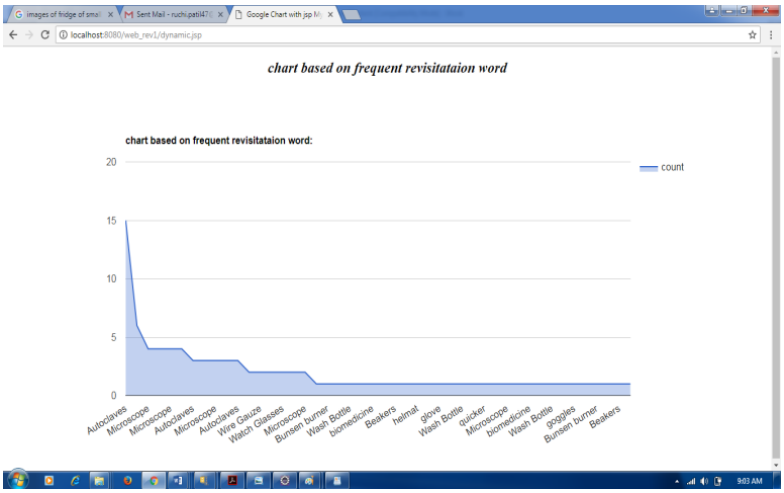

Figure 5: Area Chart

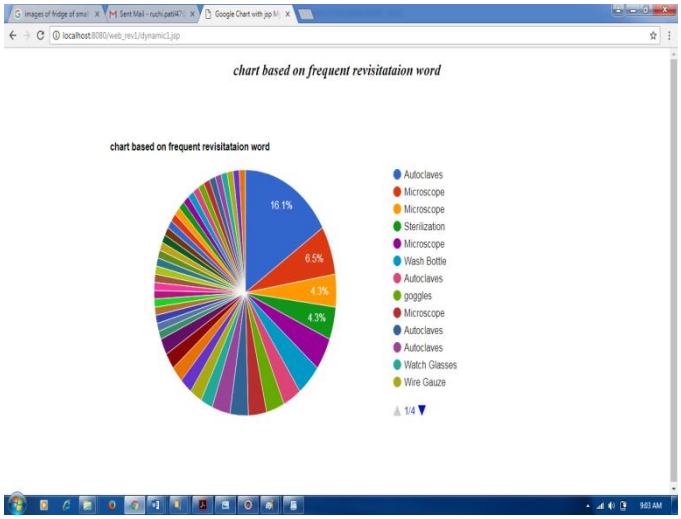

Figure 6: Pie Chart

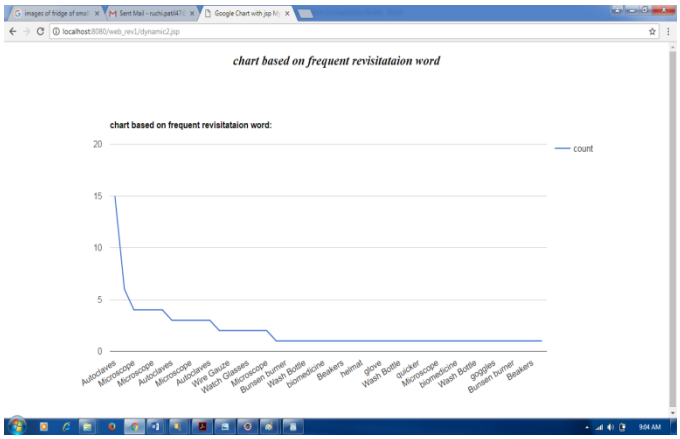

Figure 7: Line Chart

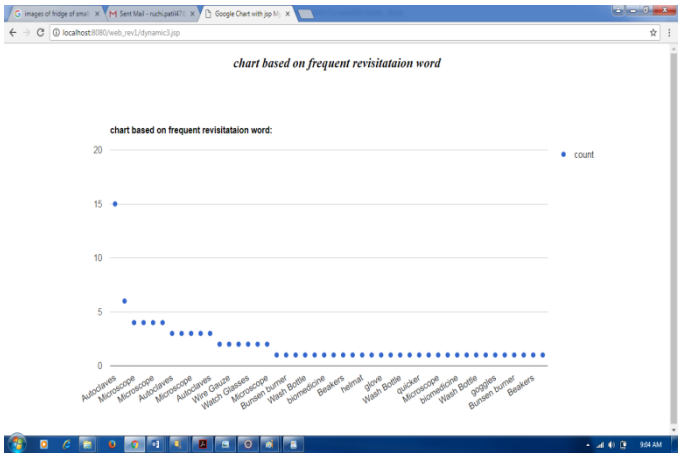

Figure 8: Scatter Chart

V. Conclusion

Drawing on the characteristics of human brain memory in organizing and exploiting episodic events and semantic words in information recall, this paper presents a web user revisitation relevance feedback using keyword for web mining.Context instances and page content are respectively organized as probabilistic context trees and probabilistic term lists, which dynamically evolve by degradation and reinforcement with relevance feedback. Our experimental results demonstrate the effectiveness and applicability of the proposed technique.

References

[1]. Naaman M, Song Y J, PaepckeA et al. Automatic organization for digital photographs with geographic coordinates. In Proc. the 4th ACM/IEEE Joint Conference on Digital Libraries, June 2004, pp.53–62.
[2]. Naaman M, Harada S, Wang Y et al. Context data in geo-referenced digital photo collections. In Proc. the 12th ACM International Conference on Multimedia, Oct. 2004, pp.196–203.
[3]. Cao L, Luo J, Kautz H et al. Annotating collections of photos using hierarchical event and scene models. In Proc. the 21st IEEE Conference on Computer Vision and Pattern Recognition, June 2008.
[4]. Joshi D, Luo J. Inferring generic activities and events from image content and bags of geo-tags. In Proc. the 7th International Conference on Content-Based Image and Video Retrieval, July 2008, pp.37–46.
[5]. Viana W, Filho J B, Gensel J et al. PhotoMap — Automatic spatiotemporal annotation for mobile photos. In Proc. the 7th Int. Symp. Web and Wireless Geographical Information Systems, Nov. 2007, pp.187-201.
[6]. Viana W, Hammiche S, Villanova-Oliver M et al. Photo context as a bag of words. In Proc. the 10th IEEE International Symposium on Multimedia, Dec. 2008, pp.310-315.
[7]. Crandall D, Felzenszwalb P, Huttenlocher D. Spatial priors for part-based recognition using statistical models. In Proc. the 18th IEEE Conference on Computer Vision and Pattern Recognition, June 2005, pp.10-17.
[8]. Dalal N, Triggs B. Histograms of oriented gradients for human detection. In Proc. the 18th IEEE Conference on Computer Vision and Pattern Recognition, June 2005, pp.886- 893.
[9]. Felzenszwalb P, McAllester D, Ramanan D. A discriminatively trained, multiscale, deformable part model. In Proc. the 21st IEEE Conference on Computer Vision and Pattern Recognition, June 2008.
[10]. Felzenszwalb P F, Huttenlocher D P. Pictorial structures for object recognition. International Journal of Computer Vision, 2005, 61(1): 55-79.
[11]. Hu J, Pei J, Tang J. How can I index my thousands of photos effectively and automatically? An unsupervised feature selection approach. In Proc. the 14th SIAM International Conference on Data Mining, Apr. 2014, pp.136-144.

[12]. Zhou W, Li H, Lu Y et al. Encoding spatial context for large-scale partial-duplicate web image retrieval. Journal of Computer Science and Technology, 2014, 29(5): 837-848.
[13]. Shotton J, Winn J, Rother C et al. Textonboost for image understanding: Multi-class object recognition and segmentation by jointly modeling texture, layout, and context. International Journal of Computer Vision, 2009, 81(1): 2-23.
[14]. Hu S, Chen T, Xu K et al. Internet visual media processing: A survey with graphics and vision applications. The Visual Computer, 2013, 29(5): 393-405.
[15]. Frome A, Singer Y, Malik J. Image retrieval and classification using local distance functions. In Proc. Neural Information Processing Systems, Dec. 2006, pp.417-424.

IOSR Journal of Computer Engineering (IOSR-JCE) is UGC approved Journal with Sl. No. 5019, Journal no. 49102.

Ruchi Patil. "Web User Revisitation Relevance Feedback Using Keyword For Web Mining." IOSR Journal of Computer Engineering (IOSR-JCE) 20.3 (2018): 1-6.