

# Hierarchical Similarity Of Context Instances

Sayyad F.S.<sup>1</sup> & P.P.Kalyankar<sup>2</sup>

<sup>1,2</sup>Department of M.E.(C.S.E.), T.P.C.T.'s College Of Engineering, Osmanabad, India.

Email:farosayyad@gmail.com, kalyankarpravin@rediffmail.com

**Abstract**— In this paper we present a methodology for how the contexts are analyzed for the context information finding in terms of web history searching in personal information management system. Context annotation and context degradation for context analysis are presented in the methodology. Hierarchical similarity finding between context attributes and context instances are implemented in this paper.

**Keywords**—context instances , context annotation ,context degradation, retention.

## I. INTRODUCTION

In nowadays peoples are searching huge amount of information on the web and saving to there system or just seen it and forgot for a small amount of time. Number of search engines are there for finding relevant information by a single search. But Sometimes we doesn't get the exact web page for the single search.

In this methodology we are trying to find how to analysed the context in the query of a user and how to annot them and how the context degradation is performed for the analysis of context instances.

Diversity of methods have been implemented for web information for access and reuse.[01/18]Number of techniques such as bookmarks, History retrieval, search engines and so on are present in the proposed system.

When user enters a query if it is similar to a previous query it obtains the current result from its cache. In this search engine we can also use textual information for the particular web search such as time of accessing , name of the author, publishing date etc. for example. Context annotation and context degradation for context analysis are presented in the methodology. Hierarchical similarity finding between context attributes and context instances are implemented in this paper.

## II. RELATED WORK

Any information that we have searched may be analysed by any web search. One of the basic theme is that information in the Personal Information Management[1]can be analysed in the context only.

While great efforts have been made to find the contextual information of any user like session, place, field of work. Context is nothing but any information that can characterize and is relevant to the interaction between a user and an application.

Techniques that are presented for refinding of information is explained in[1].In this techniques Personal Information Management System is used for the context analysis.

Context for the Metadata is presented in[12] in which properties of data contents is incorporated and indexed for personal information retrieval.

Soules and Ganger[22] developed a file search engine combining content-based search with temporal relationships between files gathered from user's file operations.

In this system we are trying to find the context attribute generation and finding the similarity between the contexts that are generated in the web history search engines.

A more general way for the analysis of context creation for the system demonstration such as activity, time and place are used for the re-finding of the user searched information[1].

One of the technique used in the[17] uses Information Finding by association. The system interface is such a designed model for the information finding by association. It consist of multiple levels of association and algorithms for collecting the association information and providing answers to real-time queries.

A comparative study of all the refinding methodologies is presented in [16] gives the basic idea of all the information finding systems for the analysis of all the methods.

Information or the history finding of the user by using the context analysis is one of the part of information refinding but information refinding is not the just information finding is explained in the[12].

How to improve a recall search of the user history is implemented in the[3].That system provides the recall capability of the user for the search history.

One of the technology used for the search engine is explained in the[14]for researching of the web pages by using the finding and refinding of the search pages. This technique is useful for the web history searching fundamentals. This system facilitate the user to search old information and answers to conflicting information goals.

Many times we are using the search engines based on context analysis we work through semantic web searching fundamentals. In such methods Contexts attributes that are inserted as a record of searching can be searched by maximum possibility of matching keywords.

Context and Page Analysis For Improved Web Search is also presented in[20].In this technique the page analysis for ranking and context analysis is also implemented.

## III. METHODOLOGY

### A. Context Memory Analysis

Context memory model is useful for context based information refinding [01].Context memory is divided into two

units: Short Term context Memory and Long term Context Memory. Short Term Context Memory is limited in capacity and lasts for few days. Long Term Context Memory is unlimited in capacity and lasts for as short as few days or as long as decades.

Long Term Context Memory consist of Two units Permanent and Evolving. Permanent unit has life long accessing experience and evolving unit will decay later.

Same methodology we will use for the implementation purpose that number of pages revisited frequencies and access logs will help us to determine the exact web page that the user need.

Number of circumstances under which we are accessing user's information is nothing but the access context. Context information may be internal or external. Such as user name, activity, agenda etc. or name of the author ,publishing date ,time etc.

### B. POSET of Context attribute and Context

Partially Ordered set of the context attributes denotes the properties for the Lattice as like Reflexivity, Transitivity and Anti-symmetry.

Whatever context we are trying to used in the methodology Are represented by 'n' context attributes such as  $(A_1 A_2 A_3 \dots A_n)$ . The Domain of each contextual attribute forms an ordered levels of abstraction which is denoted by  $Dom(A_i)$ [1].

Domain of the attributes denotes the set of attributes that are under consideration.

As the hierarchies denotes the levels of abstraction, we are using these levels for finding how the contexts are generated and how the similarity can be calculated for these context attributes.

### C. Hierarchical Similar Attributes

The Hierarchy of the context attribute is a lattice  $(H, <_h)$ , Where  $H=(h_1, h_2, h_3, \dots, h_{s-1}, All)$  having s levels corresponding to the level  $Id(1, 2, 3, \dots, s-1, s)$  and  $<_h$  is the partial order among the levels of H, such that  $(h_i <_h h_j <_h All)$  for every  $0 < i < j < s$ [1]. The edge linking two consecutive hierarchical levels  $h_i$  and  $h_{i+1}$  in H has weight in  $[0, 1]$  to express hierarchical similarity between  $h_i$  and  $h_{i+1}$ [1].

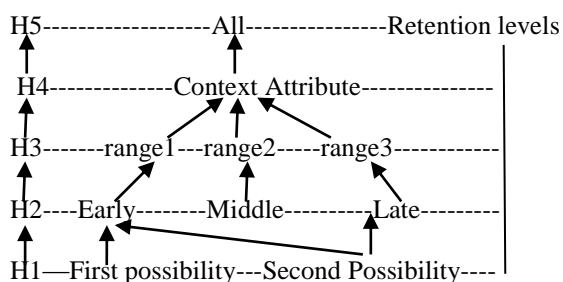


Fig.1 Contextual hierarchies of attributes.

As shown in fig.1 the possible way to represent the hierarchies of the context attribute is presented.

All these hierarchical levels also has their retention interval as the contextual attributes also get reinforced age/2 times[1].

### D. Ancestor and Descendent Of Context Attributes

The levels of the context attribute denotes the relation of their levels. When we are assuming the levels of context

attribute they gives the path from root to the destination. The levels denotes the relation  $<$  or  $>$  for the levels of abstraction.

The calculation for finding the similarity between the context attribute is subject to their distances at hierarchical levels, also their ordering distances when prompted to the same level.

Whenever we are using the search engines based on context analysis we work through semantic web searching fundamentals. In such methods Contexts attributes that are inserted as a record of searching can be searched by maximum possibility of matching keywords.

### E. Contextual similarity calculation

For finding the same information by using some of the words that are used at the time of searching is nothing but the finding of information by using the context. Many times such things are happened where the keywords are matched with the previous context and some times they are different with the input keyword information.

In this contextual attribute similarity calculation we are having two input keywords and then finding out the similarity and their behaviour for the search engines.

Given two context attribute values  $c$  and  $c'$  in the  $Domain(Attributes)$ , Consider  $c=h(A, x)$  and  $c'=h(A, y)$ . The similarity between  $x$  and  $y$  is denoted as  $SIM(A, x, y)$  and is calculated as follows.

1. If  $c = c'$ ,  $SIMILARITY(A, c, c')=1$ .
2. If  $y$  is an ancestor of  $x$  then  
 $SIMILARITY(A, c, c')=S_{k', k'+1} \cdot S_{k'+1, k'+2} \dots$
3. If  $c$  is descendent of  $y$  then  
 $SIMILARITY(A, c, c')= S_{k, k+1} \cdot S_{k+1, k+2} \dots$
4. If  $c$  and  $c'$  are two brothers of the same parent  $p$ , let  $n$  be the total number of  $p$ 's children,  $n_c$  and  $n_{c'}$  are the ordered positions of  $c$  and  $y$  among the  $p$ 's children then  
 $SIMILARITY(A, c, c')= S_{k, k+1} \cdot \frac{|n_c - n_{c'}|}{n-1} \cdot (S_{k, k+1} - S_{k, k+1}^2)$
5. If  $c$  and  $c'$  are two different hierarchical levels with a common ancestor  $p$ , let  $m=h(A, p)$ ,  
 $SIMILARITY(A, c, c')= (S_{k, k+1} \cdot S_{k+1, k+2} \cdot S_{k+2, k+3} \dots S_{m, m-1} \dots) \cdot (S_{k', k'+1} \cdot S_{k'+1, k'+2} \cdot S_{k'+2, k'+3} \dots S_{m, m-1} \dots)$

6. If one of the condition from the above steps is true we got the required information of that context attribute placed in the database of that context, and move for the similarity of context instances.

7. After calculating the combined similarity of all the attributes, calculation for context instances can be done. Context instance is an object of all its context attributes.

8. Consider context instance is denoted by  $C$  and  $C'$   
 $C=(c_1, c_2, c_3, \dots, c_n)$  and  $C'=(c'_1, c'_2, c'_3, \dots, c'_n)$

$$SIMILARITY(C, C') = \sqrt{\frac{1}{n} \sum_{i=1}^n SIMILARITY^2(A, c_i, c'_i)}$$

9. The above calculated similarity measure gives us the result of the search.

When a user wants to search a web page by using context analysis, one can insert the context of that search and the  $n$  finds the related information regarding the search. Context analysis and context annotation is mainly used in the purpose of information finding where the searched context are similar to the previous search.

#### IV. RESULTS ON SYNTHETIC DATA

When a user wants to use a search engine based on the context analysis, first consideration is for the similarity measure of context attribute as well as context instances. By using the techniques used in the proposed methodology is quite useful for finding the context similarity. If any keyword is already searched it will be inserted into the record of context attributes and then after collecting mostly matched context attribute are gathered to form a context instance.

Context are generated manually as per the user needs and can be used as a basis for the search engine methodologies. Whenever calculating the similarity between three context attribute amongst them if two context are same then the similarity measure is obtained(0.72, 0.48). and if the three context attributes are different then the similarity measure can also be calculated(0.8).

When analysis can be made on finding the similarity between two context instances with three context attributes amongst them one attribute is same, the similarity measure is also calculated (0.87).

#### V. APPLICATIONS

Number of techniques are present now a days for the re-finding of information that was previously searched because the search methods are same for the user as inserting the just keyword or the related word of search at these instances context based web searching is useful.

With the technique used above the context similarity calculation is presented as a part of web history search engine methodologies. Context similarity measure and context instance similarity measure is also presented in the above technique.

Number of methods using clustering and association of context attributes, so we can easily form the user characteristics. Decisions can be made for the user history, user behaviour, accessed information, access context etc.

As the context are annotated by means of the context attributes and context instances similarity finding, the next improvement is for the context degradation. For many methods that implements web searching by the analysis of context, first needs to annotate context and then finding similarity measure of there context attributes and context instances.

#### VI. CONCLUSION

We have used the basic fundamentals for finding the similarity of context instances and their attributes. Selection criteria is dependent on user's choice. Using any web history search engine context analysis and context annotation is the basic fundamental step.

In nowadays peoples are searching huge amount of information on the web and saving to there system or just seen it and forgot for a small amount of time. Number of search engines are there for finding relevant information by a single

search. But Sometimes we doesn't get the exact web page for the single search in this situation also context based search techniques are useful

While using context as the way of search methodologies context attributes are grouped together to form the context instance and then after clustering and association can be performed as the basis of search engines.

#### VII. FUTURE SCOPE

Context degradation is one of the important part as we are dealing with the context information finding. Context attributes value degrades along a hierarchy from specific to general. As human memory is also degraded as the age goes by, same like the unnecessary information present in the context memory snapshot is also to be removed for inserting new values of context attributes and context instances.

Another way of improving the search engines methodology is that context can be annotated automatically. This technique can be used for the better analysis of inserting and deleting the contexts for the search engines.

As context degradation gives the new context in the web searching they also needs to go with similarity measure findings of context attributes and context instances.

#### REFERENCES

- [1] Tangjian Deng,Liang Zhao,Hao Wang,Qingwei Liu and Ling Feng"Refinder:A Context Based Information Refinding System"IEEE transaction on Knowledge and Data Engineering, Vol.25,NO.9,September 2013.
- [2] M. Fuller, L. Kelly, and G.J.F. Jones, "Applying Contextual Memory Cues for Retrieval from Personal Information Archives,"Proc. Personal Information Management Workshop at CHI, 2008.
- [3] J. Hailpern, N. Jitkoff, A. Warr, R. Karahalios, K. Seseck, and N. Shkrob, "YouPivot: Improving Recall with Contextual Search," Proc. SIGCHI Conf. Human Factors in Computing Systems (CHI), 2011.
- [4] S.K. Tyler and J. Teevan, "Large Scale Query Log Analysis of Re-Finding," Proc. Third ACM Int'l Conf. Web Search and Data Mining (WSDM), 2010.
- [5] C.A.N. Soules and G.R. Ganger, "Connections: Using Context to Enhance File Search," Proc. 20th ACM Symp. Operating Systems Principles (SOSP), 2005.
- [6] Y. Chen and G. Jones, "Integrating Memory Context into Personal Information Re-Finding," Proc. Second Symp. Future Directions in Information Access, 2008.
- [7] S. Dumais, E. Cutrell, J. Cadiz, G. Jancke, R. Sarin, and D.C.Robbins, "Stuff I've Seen: A System for Personal Information Retrieval and Re-Use," Proc. 26th Ann. Int'l ACM SIGIR Conf. Research and Development in Information Retrieval (SIGIR), 2003.
- [8] S. Won, J. Jin, and J. Hong, "Contextual Web History: Using Visual and Contextual Cues to Improve Web Browser History," Proc. SIGCHI Conf. Human Factors in Computing Systems (CHI), 2009.
- [9] D. Morris, M.R. Morris, and G. Venolia, "Searchbar: A Search-

- Centric Web History for Task Resumption and Information Re-Finding,” Proc. SIGCHI Conf. Human Factors in Computing Systems (CHI), 2008.
- [10] B. MacKay, M. Kellar, and C. Watters, “An Evaluation of Landmarks for Re-Finding Information on the Web,” Proc. Extended Abstracts on Human Factors in Computing Systems (CHI’05 EA), 2005.
- [11] L. Feng, P.M. Apers, and W. Jonker, “Towards Context-Aware Data Management for Ambient Intelligence,” Proc. 15th Int’l Conf. Database and Expert Systems Applications (DEXA), 2004.
- [12] R. Capra, M. Pinney, and M.A. Perez-Quinones, “Refinding Is Not Finding Again,” technical report, Aug. 2005.
- [13] L. Tauscher and S. Greenberg, “How People Revisit Web Pages: Empirical Findings and Implications for the Design of History Systems,” Int’l J. Human Computer Studies, vol. 47, pp. 97-137, 1997.
- [14] J. Teevan, “The Re:Search Engine: Simultaneous Support for Finding and Re-Finding,” Proc. 20th Ann. ACM Symp. User Interface Software and Technology (UIST), 2007.
- [15] E. Tulving, “What is Episodic Memory?” Current Directions in Psychological Science, vol. 2, no. 3, pp. 67-70, 1993.
- [16] A.P.Nivethith,,D.KeranaHanirex, Dr.k.P.K.Kaliyamurthie  
”A Comparative Study Of Context-Based Information Refinding”  
An international journal of advanced computer technology ,2004.
- [17] Due Horng Chau, Brad Myers and Andrew Faulring ”Feldspar:A System For Finding Information By Association ” CHI 2008,Florence,Italy.
- [18] Google Web History, <http://www.google.com/history>, 2013.
- [19] Prathyusha Kanakam1, S. Mahaboob Hussain2, Dr. Sumit Gupt3, Dr. D. Surya Narayana4 “An Analysis of Exploring Information from Search Engines in Semantic Manner”.
- [20] Steve Lawrence and C.Lee Giles”Context and Page Analysis For Improved Web Serach” NEC Research Institute.
- [21] J.Z. Tsien, The Memory. Scientific Am., 2007.
- [22] E. Tulving, Elements of Episodic Memory. Oxford Univ. Press, 1983.
- [23] E. Tulving, “What is Episodic Memory?” Current Directions in Psychological Science, vol. 2, no. 3, pp. 67-70, 1993.
- [24] M. Mayer, “Web History Tools and Revisitation Support: A Survey of Existing Approaches and Directions,” Foundations and Trends in Human-Computer Interaction, vol. 2, no. 3, pp. 173-278, 2009.