International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 7s
DOI: https://doi.org/10.17762/ijritcc.v11i7s.6986
Article Received: 02 April 2023 Revised: 18 May 2023 Accepted: 28 May 2023

Perception based User Profiles for Web Personalization

Sowbhagya M P[1], Yogish H K[2], G T Raju[3]
[1]Assistant Professor, Dept. of AIML
Dayananda Sagar Academy of Technology and Management, Bengaluru
Visvesvaraya Technological University, Karnataka
Sowbhagya.mp@gmail.com
[2]Professor, Department of ISE
M S Ramaiah Institute of Technology, Bengaluru
Visvesvaraya Technological University, Karnataka
yogishhk@gmail.com
[3]Principal,
SJC Institute of Technology, Chickballapur
Visvesvaraya Technological University, Karnataka
gtraju1990@yahoo.com

Abstract—Personalized web services reduce the burden of information overload by collecting facts that match the needs of the user. An important aspect of personalized web services is the creation of user profiles that contain user information and settings. This article introduces a unique method called Perception-Based User Profiles (PUP) based on perception and browsing order, develops and updates user profiles. User profiles include perceptions and relationships, which can help guarantee that user interests are represented semantically. Second, when calculating the perception and duration of the relationship, for each site in a session, the user's browsing order is considered. Third, cognitive psychometric memory model is used to update the user profile's perceptions and relationships at the end of each session, ensuring the user profile's dynamics. The results of the tests suggest that this strategy works well for building and updating user profiles.

Keywords-User Profiles,Personalization,web Usage mining.

I. Introduction

Personalization is the use of technology to correct differences between individuals. Websites are personalized according to an individual's characteristics. Personalization denotes that changes are made based on hidden data, such as things purchased or pages visited. A solid user profiling technique is a necessary and vital part of online personalization. In the scientific literature, the term "preference" is used in a variety of ways that are related but not identical. Preferences can be defined as a person's attitude toward a group of objects. A user profile is a compilation of personal information about a certain user. As a result, a profile refers to a person's clear digital depiction of their identity. A user profile is a computerised version of a user model. The description of a person's qualities can be stored in a profile. Systems that take into account the qualities and preferences of the individuals can use this information. The term "profiler" refers to the process of creating a profile by extracting data from a set of data. User profiles serve as a kind of parameterization for request statements, capturing common variances in requests for similar systems. The document-based approach and the cognitive-based approach are the two basic methodologies for user profiling strategies. Users' clicking and browsing patterns are captured using a document-based user profiling approach. Individual users may have different functional requirements, necessitating the evaluation of different functional subsets, or they may have multiple non-functional constraints on functions, necessitating the evaluation of different functional subsets. Perception-based user profile methodologies, on the other hand, try to capture users' conceptual demands. Users' search history and browsed pages are automatically categorised into a set of subject categories. Users' preferences on the derived subject categories are used to generate user profiles. This article presents a unique method for creating and updating user profiles called Perception-Based User Profiles (PUP) to represent users' real-time preferences for web services personalized. The PUP is made up of perceptions and connections gleaned from the user's browsing history. When measuring relationship awareness and lifetime value, the sequence in which users browse each site during a session is taken into account.

II. Related Work

How to construct user profiles has gotten a lot of attention in recent years as a way to improve the qualification of Web tailored services. Some researchers utilize a technique called

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 7s
DOI: https://doi.org/10.17762/ijritcc.v11i7s.6986
***Article Received**: 02 April 2023 **Revised**: 18 May 2023 **Accepted**: 28 May 2023*

static profiling, which examines users' predictable and static characteristics. The approach of information fusion is used by Li et al to create a user profile [1] and to create a user profile; the system uses a fuzzy set based clustering algorithm [2]. Despite its simplicity, the profile becomes obsolete as the user's interests change due to a lack of dynamics. Some methods analyze URLs that the user has visited to determine the user's interests. Furthermore, it will have an impact on the qualification of personalized recommendations because it is impossible to gather the user's dynamic interests. Despite the fact that the semantic information in these user profile models allows them to properly reflect the content of the user's interested Web pages, it remains a challenge to evolve and maintain the model's adaptiveness in order to signify the consumer's interests by limited subjects. Singh et al. propose a news filtering system that models the user's preferences using an interest hierarchy based on explicit user feedback [7]. Due to the lack of semantics, dynamics, and adaptability in current research, it is difficult to adequately capture users' real-time interests that affect the quality of tailored suggestions. This research proposes a novel way for building and maintaining user profiles in order to address these issues.

III. Generating a User Profiles

Web activity by users is a constant process of information searching. The internet makes information available to users who are interested in it. As the user's surfing activity progresses, so does his or her knowledge. Increased knowledge leads to a shift in user interest, according to the information seeking theory [9]. As a result, we create a user profile based on each encounter he or she has with the Internet. In this paper, PUP is defined as a method for determining a user's real-time interest in a Web activity process. Views and connections collected from the user's Internet browsing are represented. The dynamic user profile is ensured by updating the attribute of each perception and relation.

P=<V, E> describes the PUP, where V = {v1, v2,..., vn} - finite non-empty set of vertices. Each vertex v is a pair vi= c, l>, where c denotes the semantic perception and l denotes the vertex's life span.

E = {e_1, e_2, ..., e_m} - a finite collection of non-empty edges between vertices.

Each edge e is a pair $e_i = (<v_i, v_j>, l)$ where $v_i, v_j \in V$, $<v_i, v_j>$ is yet another pair that describes the relationship between two vertices, l in a user profile, the edge's life time is defined.

The perception from Web contents that the user visited corresponds to the vertex in the specification above. A semantic relationship is represented by the edge between vertices. Both vertices and edges with a high l attribute suggest perceptions, and the relationship plays an significant role in the representation of user interest.

The PUP has two distinct characteristics, which are explained below.

A. Sematics:

The information that a user has viewed on the internet is used to construct a user profile. The perception level is used to represent it. Furthermore, Web material is used to extract relationships between perceptions. As a result, perceptions in PUP are not isolated; they are linked together, allowing for a semantically rich PUP.

B. Dynamics:

When a person interacts with Web sites, PUP changes in response to their changing interests. The active degree of each perception and relation is indicated by the attribute l of a vertex and an edge in PUP. When a user interacts with anything, the properties of some vertices and edges change to indicate the shift in user attention. As a result, the dynamics of PUP can be ensured through evolution mechanisms.

When it comes to creating PUP, there are two major concerns to consider:

i) How to extract perceptions and semantic associations from a person's first browsing session to create an initial user profile
ii) How to modernize the user profile as the user's perusing actions continue in order to establish the evolution mechanism of perceptions and semantic relations.

For the first problem, we use the perceptions and relations extraction method described in [10], which uses support value to extract perceptions from Web material. In addition, we consider the user's browsing order. To begin, each browsing session's life time of perception is calculated. Assume that the consumer browses n Web pages during the session, and that f is the frequency with which the perception occurs on those sites, then there's the initial lifespan of perception. In this page, P Linit is calculated as $PL_{init} = f / n$.

According to the Information Seeking Theory, a user's surfing activity is a process in which the user's requirements become increasingly obvious. As a result, late-session material surfing can better reflect a user's interest in that session. As a result, based on the surfing position of the Web page, the vertex life duration is altered as follows

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 7s
DOI: https://doi.org/10.17762/ijritcc.v11i7s.6986
***Article Received**: 02 April 2023 **Revised**: 18 May 2023 **Accepted**: 28 May 2023*

$$VL = PL_{init} \sum_{p=1}^{n} \frac{p}{n}$$

where VL stands for adjusted vertex life time and p stands for the browsing position of the perception-related Web pages. It's been revealed that perceptions displayed later in the browsing process play a bigger influence in portraying user interest. Perceptions with values greater than the threshold can be chosen for PUP based on the computed value.

Based on Point Wise Mutual Information (PMI)[11], semantic relationships between perceptions can be retrieved in PUP. The life duration of the RLinit relationship between the perceptions ci and cj can be calculated as follows:

$$RL_{init}(c_i, c_j) = \log \frac{N * vf(c_i \cap c_j)}{vf(c_i) * vf(c_j)} / \log N$$

where vf(cicj) is the combined page frequency of the ci and cj, and log N is a normalization factor that keeps the value between 0 and 1. Adjusted life time RL(ci, cj) of semantic relation is determined similarly to perception life time.

$$RL(c_i, c_j) = RL_{init}(c_i, c_j) \sum_{q=1}^{n} \frac{q}{n}$$

The surfing position of Web sites involving both the senses ci and cj is denoted by q.

Relations with values greater than the threshold can be chosen for PUP based on the computed values. As a result, the first PUP file for the customer's first browsing term is created.

IV. PROFILE OF THE USER UPDATING

PUP should be updated based on fresh surfing contents because the user profile reflects the individual's interests, which can be gleaned through browsing history. Following the completion of a new browsing session, the same procedure used to create the initial PUP file can be used to create the appropriate session semantics. Then, when it comes to updating PUP, it all boils down to figuring out how to merge the new session semantics file into the prior PUP while also updating the old vertices and edges.

Finally, complete content and organizational editing before formatting. Please take note of the following items when proofreading spelling and grammar:

A. In PUP, Vertices and Edges' Lifetimes are being updateds

The updating of PUP is dependent on the updating of vertex and edge life times. The life time of newly linked vertices and edges in PUP is rather long, while the life time of vertices and edges that will be erased from the PUP file is short. The following principles are considered when updating life time:

- The vertices and edges that reflect the most recently widely visited contents should have a long life, and contents that the user has previously browsed but not recently are excluded from the user profile.
- A user-interested vertex or edge will stay in PUP longer than those that are unable to convey the user's preference. To put it another way, this type of vertex and edge should have a long lifespan.
- PUP is updated based on its life time. If the life duration of one vertex or edge is less than the threshold, it is not adequate for expressing the user's interests and will be removed from PUP.

This type of vertex should be kept in the modified PUP file since it shows that the user is still interested in perception-related things in the new session. The average value of the two is computed in this research. Let lup signify the new session semantics files updated life attribute value and ls denote the old session semantics file's life attribute value, then

$$l_{up} = {(l_{up} + l_s)}/{2}$$

Aside from the types of vertices outlined above, there is another type of vertices that was present in previous PUP but is absent from the current session semantics file. They may not appear in the user's new browsing session, however, this does not imply that they are unrelated to the user's interests. The user profile's functionality also includes the ability to reflect the individual's most recent interests. As a result, the fact that this vertex did not appear in a single session does not rule out the possibility that it is implicated in PUP.

We adopt the memory capacity model SIMPLE (Scale Invariant Memory, Perception, and Learning) from paper [12] to describe this problem. For vertices in an old PUP that aren't implicated in the new session semantics file, the probability of forgetting an item is used as an adjustment factor. According to SIMPLE, the probability of recalling a particular item I is proportional to the psychological distance between item I and the item j before it, and the probability p is computed as follows:

$$p = e^{\delta * dist}$$

where dist is the psychological distance between item I and item j, and is the psychological distance between item i and item j, the adjustment factor is. The temporal and spatial distance between two items is referred to as psychological distance.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 7s
DOI: https://doi.org/10.17762/ijritcc.v11i7s.6986
***Article Received**: 02 April 2023 **Revised**: 18 May 2023 **Accepted**: 28 May 2023*

As an extension, psychological distance is calculated in this work based on the two factors listed below.

- In terms of semantics, the old PUP file and the new session semantics file differ. Obviously, the semantic differences between the old PUP file and the new session semantics file. Here, we compute the difference degree using the COSINE value as sd=1-cosine.
- The degree of temporal divergence between the vertex's latest emergence and the updating moment. The number of sessions between its last arising and the updating moment ,it is used to calculate the temporal difference degree: td = (n - i)/ (n-1), Where n denotes the total number of browser sessions for the user. When updating PUP for the 16th session, for example, if a perception arises in the 2nd, 3rd, 6th, 9th, and 10th sessions, (16-10)/(16-1)=0.4 is used to get the temporal distance.

The weighted sum of semantic and temporal difference degrees is used to calculate psychological distance dist for a vertex, taking into account the two elements mentioned above.

dist= w∗sd+ (1− w)∗td

By computing the likelihood of recalling vertex i from the present PUP file but not from the new The adjusted life period of a vertex that is involved in the current PUP file but not in the new session semantics file can be approximated using the session semantics file.

$$l_{up} = (1 - p) * l_{up}$$

The updated life time of a node is observed to decrease in this instance. If this vertex does not appear in the next several sessions, its life will be cut short, and it will be removed from PUP. In the same way, existing file edges are separated into two types:

- It appear in both the new session semantics file and the previous (existing) UP-CR file.
- appear in the previous UP-CR file but not in the new session semantics file.

An update method similar to that used for vertices is employed for these edges.

B. Adding and removing vertices and edges from a PUP

The new session semantics file will have some new vertices and edges after a fresh browsing session is completed. They are obviously very relevant to the user's interests. As a result, they'll be included in the existing PUP file. The life durations of vertices and edges have been included to the PUP file in the new session semantics file. Figure 1 shows how to modify an existing PUP file by adding additional vertices and edges. As a user browses more Web content, his or her interests are likely to shift, which implies the vertices and edges in the old PUP file indicating his or her former interest should be erased. In this work, the premise of eliminating vertices and edges is that if a perception does not appear in numerous neighbouring sessions, then this vertex and its edges with other vertices cannot represent the user's interest. The life duration of vertices and edges in the previous PUP file is altered when a new session semantics file is produced, and using the method mentioned above, some vertices and edges are added to the existing PUP. Then there are n files with session semantics from recent browsing sessions will be examined. A vertex will be removed from an existing PUP file if it does not present in the n session semantics files. All edges that are associated with it will also be removed. A vertex or an edge may also be removed from PUP in some cases. It will not be reserved in the PUP file if its life expectancy is less than the deletion threshold. PUP has a relatively small amount of vertices and edges. On the one hand, when developing a session semantics file, high life time values for vertex and edge are chosen. When updating an existing PUP file, however, certain vertices and edges will be removed. As a result, the size of vertices and edges can be adjusted.

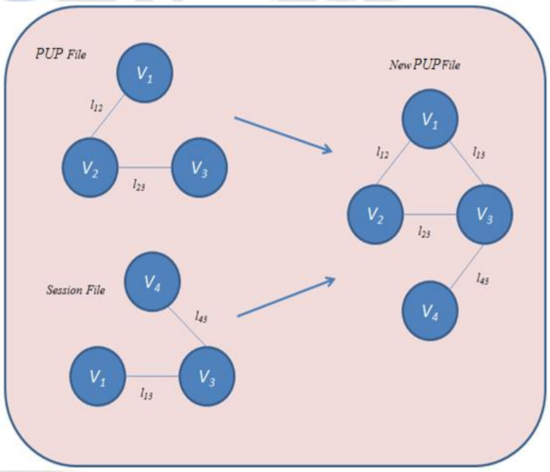

Figure 1. Adding new vertices and edges to a PUP file

V. EXPERIMENTAL RESULTS AND DISCUSSION

We used a trace of a user's browsing contents in Web of Science's academic database to demonstrate the proposed technique. To begin based on the user's surfing sequence and contents, a customer profile for a period in which he read five publications is produced as shown in Figure 2. "Collaborative filtering" is the query key term for this session. The relationship threshold is set at 0 while the perception threshold is set to 0.2. The user then went through two more sessions, each of which he read four papers. As a result, the user profile files are

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 7s
DOI: https://doi.org/10.17762/ijritcc.v11i7s.6986
***Article Received**: 02 April 2023 **Revised**: 18 May 2023 **Accepted**: 28 May 2023*

modified, as seen in Figure 3. The number of vertices and edges does not appear to expand rapidly during the updating process. According to the semantics of vertices, The validity of the user profile generation and updating approach presented in this work is demonstrated by user interests in collaborative filtering, user interest, and tailored search.

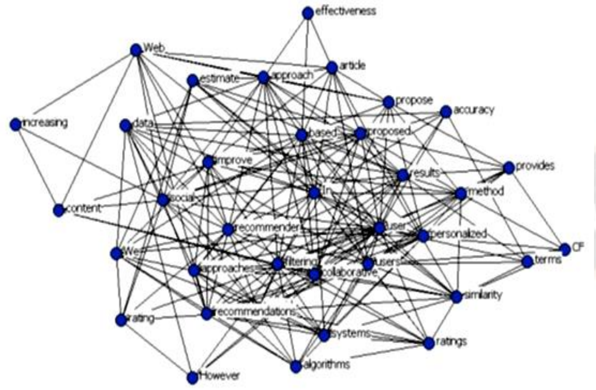

Figure 2. PUP of the 1st session

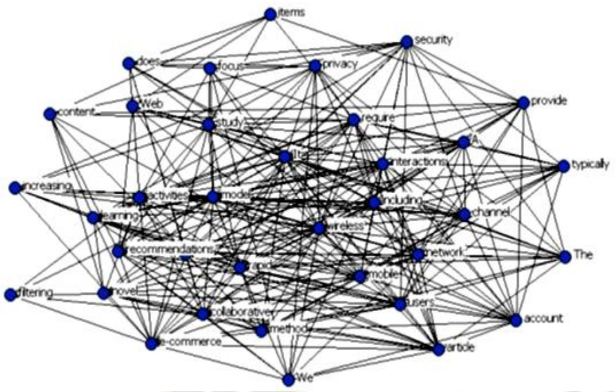

Figure 3. Updated PUP after the 2nd Session

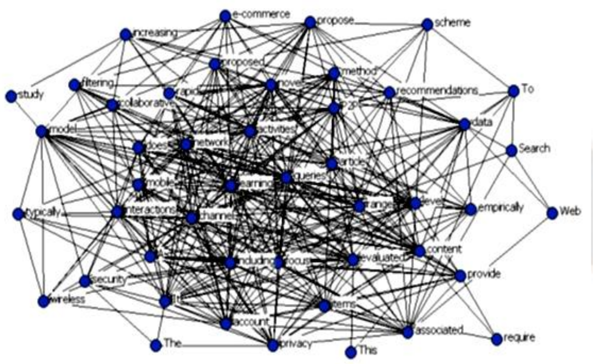

Figure 4. Updated PUP after the 3rd Session

VI. CONCLUSION

In the field of personalized services, the process of creating and updating user profiles is gaining popularity. We've described a new way for creating and updating user profiles based on perception and relationship in this article. When constructing a user's profile, the sequence in which they browsed is taken into account, ensuring that the user's interests are accurately captured. Memory capacity modeling was also addressed while updating user profiles to account for dynamic changes in user preferences. The proposed method's efficiency is demonstrated by experimental data. This is seen to hold a lot of potential in the field of personalized services.

REFERENCES

[1] X. Li, S. K.Chang. A Personalized E-Learning System Based on User Profile Constructed Using Information Fusion. In Proceedings of the 11th International Conference on Distributed Multimedia Systems, Banff, Canada,2005, pp.109-114

[2] Ushasree, D., Krishna, A. V. P. ., Rao, C. . M. ., & Parameswari, D. V. L. . (2023). SPE: Ensemble Hybrid Machine Learning Model for Efficient Diagnosis of Brain Stroke towards Clinical Decision Support System (CDSS). International Journal of Intelligent Systems and Applications in Engineering, 11(1), 339–347. Retrieved from https://ijisae.org/index.php/IJISAE/article/view/2544

[3] F. Lu, X. Li, Q. T. Liu. Research on Personalized E-Learning System Using Fuzzy Set Based Clustering Algorithm. Lecture Notes in Computer Science. Springer Berlin, Heidelberg 2007.

[4] B.Mobasher, H.Dai, T. Luo, Y.Q. Sun. Integrating Web Usage and Content Mining for More Effective Personalization. In Proceedings of the 1st International Conference on Electronic Commerce and Web Technologies, London, UK, 2000, pp. 165 - 176.

[5] F. Qiu, J. Cho. Automatic Identification of User Interest for Personalized Search. In Proceedings of the 15thinternational conference on World Wide Web, Edinburgh Scotland, 2006, pp. 727 – 736.

[6] A. Sieg, B. Mobasher, R.Burke. Ontological User Profiles for Representing Context in Web Search. InProceedings of the 2007 IEEE/WIC/ACM International Conferences on Web Intelligence and Intelligent Agent Technology, Silicon Valley, USA, 2007 pp. 91- 94.

[7] Ms. Pooja Sahu. (2015). Automatic Speech Recognition in Mobile Customer Care Service. International Journal of New Practices in Management and Engineering, 4(01), 07 - 11. Retrieved from http://ijnpme.org/index.php/IJNPME/article/view/34

[8] L. Chen, K. Sycara. WebMate: Personal Agent for Browsing and Searching. In Proceedings of the 2ndInternational Conference on Autonomous Agents, Minneapolis, USA, 1998, pp. 132-139.

[9] S. Singh, M. Sarabdeep, J. Duffy. An Adaptive User Profile for Filtering News Based on a User Interest Hierarchy.In Grove, Andrew, Eds. Proceedings 69th Annual Meeting of the American Society for Information Science and Technology, Austin, USA, 2006.

[10] J. Carroll, M. Rosson. The Paradox of the Active User. Interfacing Thought: Cognitive Aspects of Human- Computer Interaction, 1987, pp.80 – 111.

[11] Rossi, G., Nowak, K., Nielsen, M., García, A., & Silva, J. Enhancing Collaborative Learning in Engineering Education with Machine Learning. Kuwait Journal of Machine Learning, 1(2). Retrieved from http://kuwaitjournals.com/index.php/kjml/article/view/119

[12] P. Hansen, K.Järvelin. The information seeking and retrieval process at the Swedish patent and registration office. Moving from lab-based to real life work-task environment. Proceedings of the ACM SIGIR 2000 Workshop on Patent Retrieval, 2000, pp.44–53.

[13] J. Yu, et al. Building Search Context with Sliding Window for Information Seeking. Proceedings of 2011 IEEE 3rd International Conference on Computer Research and Development, 274-277.

[14] Shanthi, D. N. ., & J, S. . (2021). Machine Learning Architecture in Soft Sensor for Manufacturing Control and Monitoring System Based on Data Classification. Research Journal of Computer Systems and Engineering, 2(2), 01:05. Retrieved from https://technicaljournals.org/RJCSE/index.php/journal/article/view/24

[15] D. Bollegala, et al. Measuring Semantic Similarity between Words Using Web Search Engines, In Proceedings of the 16th international conference on World Wide Web, 2007, pp.757–786.

[16] I. Neath, G. D. A. Brown. SIMPLE: Further Applications of a Local Distinctiveness Model of Memory. Psychology of Learning and Motivation, Vol.46 . pp. 201-243. ISSN 0079-7421.

[17] Juan Lopez, Machine Learning-based Recommender Systems for E-commerce , Machine Learning Applications Conference Proceedings, Vol 2 2022.