Journal of Theoretical and Applied Information Technology
31st July 2020. Vol.98. No 14

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

AN ONTOLOGY-BASED PROFILING METHOD FOR ACCURATE WEB PERSONALIZATION SYSTEMS

[1]ANAS EL-ANSARI, [2]ABDERRAHIM BENI-HSSANE, [3]MOSTAFA SAADI

[1,2] LAROSERI laboratory, Computer Science Department, Sciences Faculty,

Chouaib Doukkali University, El Jadida, Morocco

[3] Univ Sultan Moulay Slimane, LaSTI laboratory, ENSA Khouribga,

B.P 77 Khouribga, Morocco

E-mail : [1]anas.elansari@gmail.com, [2]abenihssane@yahoo.fr, [3]saadi_mo@yahoo.fr

ABSTRACT

In recent years, the huge development in information technology led to a data explosion on the web, motivating the need for powerful and efficient strategies for information retrieval. Personalized Web systems are an example used to enhance the user experience by offering tailor-made services according to his profile. Building accurate profiles representing the reel user's interests that can change in time is the major ingredient for an efficient personalization system. This work presents our approach for generating accurate and dynamic user profiles implicitly by tracking and capturing the user's interests and preferences. Moreover, we investigate techniques to improve the profiles' accuracy; through accumulating more browsing data from multiple sources, distinguishing the most relevant concepts, and also identifying the number of ontology levels in the concepts' hierarchy needed to accurately represent each user's reel interests and preferences. Exploiting users' feedback, results prove feasibility and accuracy of the generated profiles.

Keywords: *Web personalization system, User profile, accuracy, Ontology.*

1. INTRODUCTION

The unprecedented growth of data available on the web has far surpassed traditional processing capabilities [3], preventing users from obtaining the desired service, information, or product.

This problem highlights an urgent demand for efficient personalized web systems that simplify content discovery and information access for the user considering his preferences and needs.

An example of Web personalization systems that become popular lately is the personalized search engine. These systems present personalized answers to users about desired items, services, or information. While traditional search engines return a massive list of web pages, that might contain the answer to the user query. Without a doubt, in a traditional search engine, the same set of search results is returned for all users typing the same query regardless of their intentions and needs.

There are many applications for personalization systems that support users with their learning, shopping, entertainment, traveling, or other personal requirements. Agoda.com, for example, recommends personalized tourist hotels and places for users based on their location and preferences. Another example is Amazon.com which recommends products of interest to its consumers. YouTube.com also offers personalized video recommendations.

Although creating an accurate user profile that represents the user's reel interests that can change in time is a big challenge in the personalized systems field [8]. Therefore, it is essential to focus on the user to know his interests and needs following any interest change to build efficient Web personalization systems.

In this work, the main objective of our research is to implicitly generate an accurate and dynamic user profile based on a reference ontology. Mainly we use implicit methods for creating profiles with the capacity to adjust in time, reflecting any change in user interests.

The remainder of this paper is arranged as follows; Section 2 addresses related studies. Section 3, presents the proposed method for creating accurate user profiles implicitly. Section 4 presents the user profile structure. In section 5 we discuss the experimentation results. Finally, we conclude this paper in Section 6.

2. RELATED WORK

This work relates to previous studies published in the field of information systems in general, and personalized web applications particularly. Personalized web applications are used widely to help people; purchase desired products (eBay [17], Amazon [11]), browse news articles (UP-TreeRec [27], NPA [5], SIREN [10]), find research papers (Personalized reading system [28], Pique [14]), improve search results (Persona [25], Hide-n-Seek [23]) or even to combine some of these tasks (Syskill and Webert [19], Basar [26]). These systems differ based on the sources of data collected for building the user profile; Most collect and analyze visited Web pages.

Other sources also have been used for that purpose, such as user queries in search engines and click-through data in [22], Syskill and Webert [19], and Persona [25] or bookmarks in Basar [26]. To represent user profiles, Persona [25] creates hierarchically-arranged collections of ontology concepts, while authors in [19] use a concepts' list of interest, and Basar [26] generates a bookmark-like web pages' list.

The profiles are constructed using a variety of techniques; mainly the well-known VSM (Vector Space Model) [12] [20] which is based on the notion of similarity, probabilistic model [29], genetic algorithms [16], or clustering [18]. Since the user profile may contain some noise in the form of irrelevant concepts, rating and filtering algorithms can be used to enhance the profile's accuracy. Some systems seek user feedback for this, e.g., Syskill and Webert [19], Basar [26], Persona [25]. Others adapt autonomously.

PVA in reference [4] provides personalized news articles classification. It generates a concept hierarchy-based user-profiles implicitly, from the visited pages. Building the profile by classifying the visited pages using the VSM, then adjusted by merging/splitting the concepts in the initial profile.

Another system in reference [9] learns user interest from browsed pages, and clusters the words extracted from those pages to identify concepts of interest in the form of a tree in which the leaf node is considered as a specific or short-term interest while the parent node is considered more general or long-term interest.

Authors in reference [12] cluster browsed pages and bookmarks using the VSM, and track interest shifts following any variation in the profile.

In our approach, we build an initial profile (session profile) by analyzing user browsing data (Visited URL, time spent, date, Etc.) with a browser add-on called Meetimer [15] that we adapted to implicitly track user's browsing behavior including the visited pages URLs, timestamps, and durations. To reduce processing time, we use the pages RSS feed (if available). Each visited Web page is then transformed into a keyword vector and the previously mentioned vector space model [2, 12] is employed to classify each web page into the most similar concept in the ODP reference ontology.

The generated user profile is represented by the selected concepts with their associated pages' numbers. To increase the initial profile accuracy, we exploit other internal user data sources; in particular bookmarks, search queries, and click-through data. Experimental results prove that all browsing data sources that we combine in our method improve the user profile accuracy.

3. THE PROPOSED METHOD

To help people find the right information, personalized systems have emerged. Mainly, those systems are based on two popular approaches or techniques: The first one is called Content-Based (CB), which focuses on collecting user data such as purchased products, viewed items, ratings, etc. to compute the user-item similarities, then suggest for each user new recommendations. The second named Collaborative-Filtering (CF) tracks user's browsing activities implicitly to build a profile that reflects his preferences and interests, then proposes the recommendations based on the generated profile.

This section presents our method (based on the CF technique) in which we focus, at first, on creating ontology-based profiles by tracking user's browsing activities implicitly, and second on how to improve the accuracy of the obtained user profile.

3.1 Building initial user profile

To build profiles, we collect user browsing data, mainly the Web pages visited by the user. Then we classify each page to the most similar and related concept from a reference ontology (ODP). This process consists of 3 main phases:

1) Preparing the Reference Ontology.

2) Collecting user browsing data.

3) Classifying the browsing data

3.1.1 Preparing the Reference Ontology

An ontology is the specification of a set of concepts and the relations between them. [7]. Ontologies are used for building user profiles to address the cold-start issue [13]. Personalized systems usually perform badly until collecting enough relevant user data. Adopting ontologies as a foundation for building user profiles helps in mitigating this dilemma since the collected data is aggregated with existing ontology concepts.

In our work, the ODP ontology [1] (Open Directory Project) is used as the reference ontology. In more detail, the process of preparing this ontology consists of creating, for each concept, a vector containing important related terms and their weights. Each ontology concept is related to a number of Web pages grouped and saved as training data, creating collections of super documents (one per concept). The later is then processed to remove stop-words and stemmed with porter stemming [6] algorithm to clear away common suffixes.

Following this process, we compute and store the concept vector containing the list of related key-terms and the weight of each term. Hence, each concept is represented by an n-dimensional vector (n : the number of relevant and unique terms associated with the concept).

The term weight in the concept vector is calculated using TFxIDF. In more details, UW_{ij} (Unnormalized Weight of term i in concept j) is computed as follows:

$$uw_{ij} = tf_{ij} \times idf_i \tag{1}$$

Where:
tfij = number of occurrences of ti in sdj
sdj = the super document used for training concept j

$$idf_i = Log\left(\frac{x}{y}\right) \tag{2}$$

Where:
x = Total number of documents in the collection
y = Total number of documents in the collection that contain the term t_i

The final normalized weight w_{ij} for term i in concept j, is calculated as follows:

$$w_{ij} = \frac{uw_{ij}}{\sqrt{\sum_{t_i \in d_j} (uw_{ij})^2}} \tag{3}$$

3.1.2 Collecting user browsing data

After a browsing session, we collect and analyze user's browsing data; visited web pages, bookmarks, search engines' queries, and the associated Click-through Data.

- **Visited Web pages:**

Using the visited page's URL, the system extracts an RSS file (if it exists) or the HTML file. The reason why we first scan for the RSS file is because this feed files contain a less noisy and simple version of the visited page, which can help in reducing the processing time. The system extracts these documents filtering small files that are too short to contain any relevant information (Size < 1 KB) and those on which the user spent little time (< 5 seconds), like pop-ups, irrelevant pages, or the silently redirected ones. The browsed pages are then analyzed and processed to get Web pages' vectors.

- **Search queries:**

The search query is the text the user types in a search engine to look for a piece of information satisfying his needs. Expressed with a set of keywords, search queries are usually different from other standard query languages (governed by rigorous syntax rules) and can be written in natural language.

In each browsing session, we collect the user search data in the form of a set Q (q1, q2, ..., qm) of user queries, each one is associated with a click-through document set D (d1, d2, ..., dn). The click-through documents' set D for a query q is a set of documents chosen by the user from the returned results' list. Each pair (qi, Di) is then processed the same as the browsed pages to produce a query vector.

- **Bookmarks:**

Links to Web pages stored by the user to visit later. We use Bookmarked pages mainly to increment the corresponding concepts' weights in the user profile

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

3.1.3 Classifying the browsing data

We use the same process for the collected user browsing data as described in section 3.1.1 (ODP reference ontology preparation). The browsed data vectors are generated with the same steps as the concepts' vectors.

The number of the selected terms from the web page is not fixed, therefore the highest weighted 20 terms are used to represent the page's content.

The classification process involves comparing the browsing data vector with each concept's vector from the ODP ontology (created and stored previously) by computing the cosine similarity.

In more details, the similarity of a concept c_j with a browsed page p_k is computed with the following formula:

$$Similarity(c_j\,,p_k) = \sum_{i=0}^{n} w_{ij} \times p_{ik} \qquad (4)$$

Where:
$\boldsymbol{n}$: number of unique terms in the vocabulary
$\boldsymbol{w_{ij}}$: the normalized weight of term i in concept j
$\boldsymbol{p_{ik}}$: the unnormalized weight of term i in page k

After computing the user browsing data vectors' similarity with all the concepts vectors from the ODP ontology, the results are analyzed to classify every browsing data vector to the top-matching ontology concept associated with its weight (similarity value).

This process is used for each visited page, each query, bookmarked page, etc. Then, for each ontology concept, the weight is the sum of all the associated browsing data vectors. For each ontology concept, the weight is calculated by summing all its sub-concepts' weights.

As a result of this process, our system creates an initial profile (in Figure.1) containing all concepts with weight > 0. This first version of the user profile is then subject to an optimization process to enhance its accuracy.

3.2 User profile optimization

In order to improve the user profile's accuracy, we followed the subsequent measures:

1. Examining the user profiles' stability and identifying the browsing data volume needed to achieve profile stability.
2. Rank-ordering the concepts of the profile to prune irrelevant ones to produce a more precise and accurate profile.
3. Determine the ontology levels' number that is enough/needed to build an accurate profile.
4. Exploit the user's feedback.

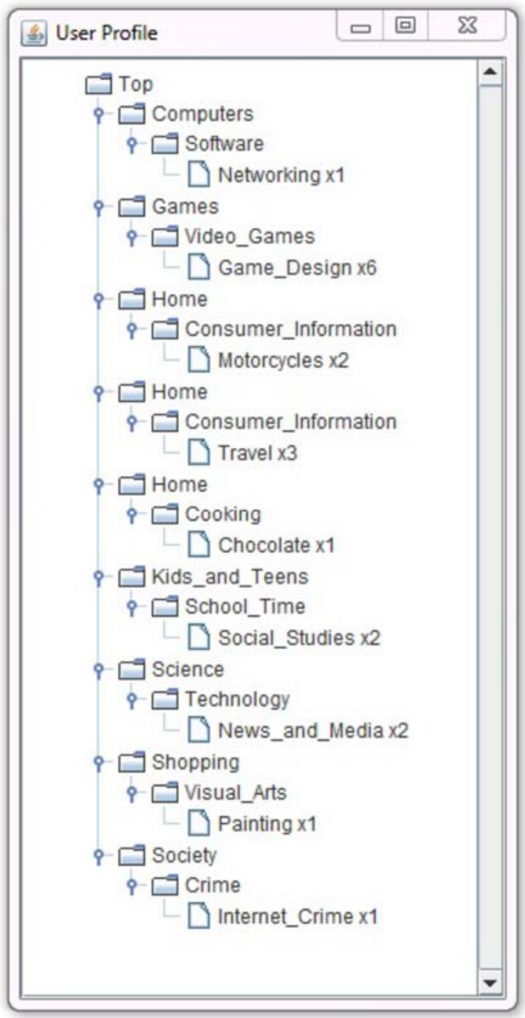

Figure 1: Screen shot of an initial user profile

3.2.1 Profile stability

Once a browsing data vector is classified, this either adds a concept to the profile or increases an existing concept's weight. We anticipate that even though the number of ontology concepts in the user profile increase in time, eventually the key-concepts in the user profile will become relatively stable, reflecting the major user interests.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

To determine the browsing data volume needed to achieve a reasonably stable profile, we examined metrics based on the time and the browsing data amount (Fig.2).

In both cases, to observe if (and when) the profile becomes stable, we measured the concepts' number in the user profile and the similarity amongst the top-ranked 50% of the concepts over time.

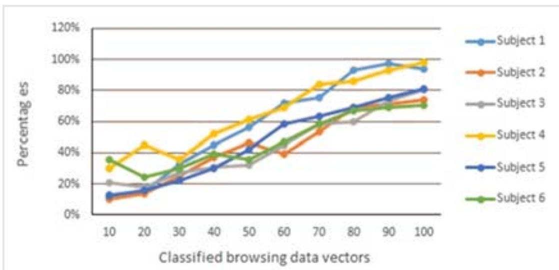

Figure 2: Similarity between top 50% of the concepts versus #of classified Web pages

Users have various browsing habits and therefore the concepts' numbers in the profiles vary greatly. This number continuously raises and shows no convergence in the short-term.

User profiles also showed no convergence as the concepts' numbers in the profiles were plotted on the amount of browsing data collected. Though, when only the top-ranked 50% of the concepts (ranked by the associated pages number) were considered, the user profile behavior showed low stability.

3.2.2 Concepts Ranking/Ordering

We can rank concepts in user profiles either by weights or by the number of browsed data vectors associated with each one. We evaluated which ranking method produces more accurate profiles by computing the F-measure. Besides, for each method, we calculated the average ranks of irrelevant concepts.

The point of ranking/ordering concepts by importance was to eliminate irrelevant ones and produce an accurate profile. Next, to evaluate the user profile accuracy, we compute the F-measure value based on different concepts' amount kept on various cut-offs. We determine the cut-off value producing the most accurate user profile based the amount with the best F-measure value.

3.2.3 Ontology levels

To determine the effect of the number of levels used from the ODP ontology on the user profile accuracy, we have experimented building profiles with one level first, then with two, and three levels from the ODP ontology concept-hierarchy. We have calculated, for each profile, the corresponding precision from the user's judgments of the profile relevance.

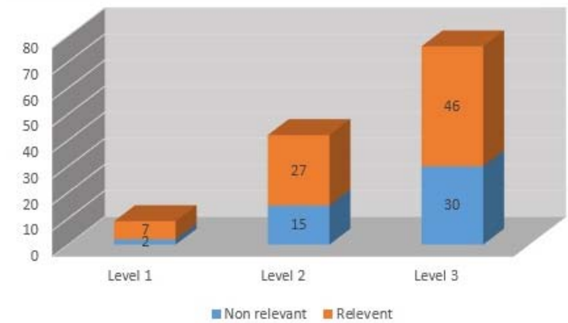

Figure 1: Number of relevant and non-relevant concepts in the user profiles versus the number of levels used in the ontology

When showing concepts from the 1st level only in the reference ontology, most user profiles contained a maximum of nine concepts, representing users' broad interest in Food, Sports, Arts, Business, etc. Though there is little information about the user's specific interests and most profiles were similar while users' reel preferences are different.

As described in (Fig.3), the number of concepts in the profile grows as we raise the levels number, allowing for 42 concepts per user on average when using two levels, and 76 concepts per profile if three levels are used. Consequently, the profile specificity increases as the depth increase.

3.2.4 Exploiting user’s feedback

To ensure the profile's accuracy, a first profile version is presented to all users through a GUI (graphical user interface). From this interface, the user can review the classified concepts on his profile and perform several operations; For instance, he can add a new concept from the ontology list of concepts, mark a concept in his profile as irrelevant, a short-term, or long-term interest. He can also increase/decrease a concept’s weight value, or even move a concept to the archive layer.

From a privacy protection perspective, users can also mark concepts as private or permanently delete them. Users are required to give their feedback on the first version of the profile and at any moment after.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Feedback is optional for each user and helps increase profile accuracy. It also helps in preserving the user's privacy and build trust. The collected feedback data is used later in the evaluation process.

4. IMPROVED PROFILE STRUCTURE

To build a user profile, we have to track and analyze the user browsing behavior, this process is initiated after every browsing session.

In our model, the generated profile is stored on the user's computer (client-side). For reasons discussed later, we create a profile separated on 3 layers or sub-profiles:

- **Short-term.**
- **Long-term.**
- **The archive.**

After each user browsing session, a set of concepts is processed and prepared to be classified to the short-term or long-term layer. A user can suddenly become interested in a subject or a topic but once he loses this interest, he simply disregards it. For example, when a Football competition like the World Cup starts, most sports fans would start following this sport. Once this event ends, they would likely turn their focus to other sports, or new events.

The short-term layer includes user recent interests. A way to discover the short-term interests is by selecting new concepts with a weight higher than a pre-defining a threshold as short-term interests. Eventually, some short-term interests may become long-term ones, and will be placed in the long-term layer.

Long-term interests are more stable than the short-term ones since they appear regularly in a long period in the user profile. For example, programmers might have an interest in programming languages that appears more regularly in their profiles as a long-term interest. Thus, the short-term sub-profile reflects changing user interests and needs while the stable ones are listed in the long-term layer.

The archive contains topics that are not of interest to the user anymore. A concept, that is gradually losing weight value (importance), is eventually placed in the archive layer based on a pre-defined weight threshold.

We made use of this profile structure for the subsequent reasons:

- The layered profile helps the system adjust to the changes in user interests.
- With the short-term and long-term layers, we can separate constant from occasional interests.

5. EXPERIMENTAL RESULTS AND DISCUSSION

5.1 Experiments and Data Collection

After the implementation stage, we conducted initial experiments on 2 system versions with the participation of 10 volunteers to measure the user profile accuracy.

In a 1st version, we used only the browsed Web pages to generate user profiles. While we combined all browsing data sources (in Section.3.1.2) in the 2nd one, to improve the profile's accuracy. We asked each user to browse freely in 3 separate sessions, then provide his feedback on the produced profile using the GUI we discussed in section 3.2.4. The collected feedback data is then prepared and analyzed to evaluate the accuracy of the produced user profiles and also the satisfaction of the participants

Since the whole process of profile construction is relying on information retrieval, measuring the profile's accuracy can be conducted with metrics used for evaluating an information retrieval system [21].

For that reason, we have calculated the F-measure, recall, and precision metrics considering all concepts in every subject's profile. The following tables and graphs represent those metrics calculated based on each user's feedback data:

Table 1: Collected Feedback Data For The 1st Version

Subjects	User's Feedback (1st version)			Precision	Recall	F-measure
	X	Y	Z			
S1	12	10	12	0,833	0,833	0,833
S2	10	10	11	1,000	0,909	0,952
S3	5	4	7	0,800	0,571	0,667
S4	14	11	13	0,786	0,846	0,815
S5	9	8	9	0,889	0,889	0,889
S6	12	12	13	1,000	0,923	0,960
S7	11	9	10	0,818	0,900	0,857

Journal of Theoretical and Applied Information Technology
31st July 2020. Vol.98. No 14

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

S8	6	5	6	0,833	0,833	0,833
S9	18	15	17	0,833	0,882	0,857
S10	15	13	16	0,867	0,813	0,839

Where:

X: The total number of concepts in the profile

Y: The total number of relevant concepts in the profile (detected by the system)

Z: Total number of relevant concept (given by the user)

Table 2: Collected feedback data for the 2nd version

Subjects	**User's Feedback (2nd version)**			**Precision**	**Recall**	**F-measure**
	X	**Y**	**Z**			
S1	12	10	12	0,875	1,000	0,933
S2	10	10	11	0,917	0,917	0,917
S3	5	4	7	1,000	0,889	0,941
S4	14	11	13	1,000	0,923	0,960
S5	9	8	9	0,900	1,000	0,947
S6	12	12	13	0,857	0,923	0,889
S7	11	9	10	0,933	0,933	0,933
S8	6	5	6	0,800	0,800	0,800
S9	18	15	17	0,850	0,895	0,872
S10	15	13	16	0,944	0,944	0,944

5.2 Results Discussion

To compare and analyze the previous results, the following graphs represent, for the tree metrics (Precision, Recall, and F-measure), a comparison between the first and second versions tested in the experiment section.

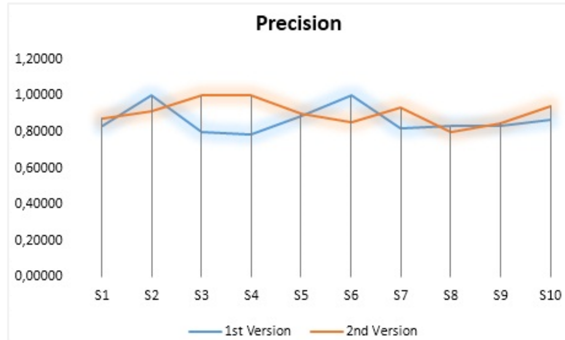

Figure 4: Precision calculated for all subjects

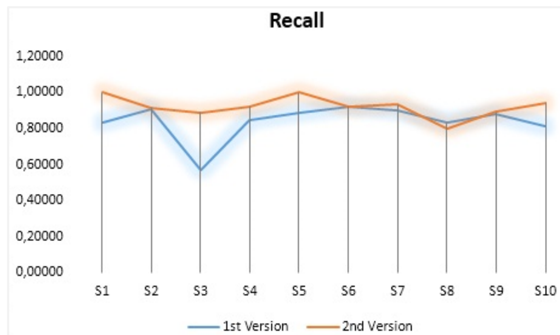

Figure 5: Recall calculated for all subjects

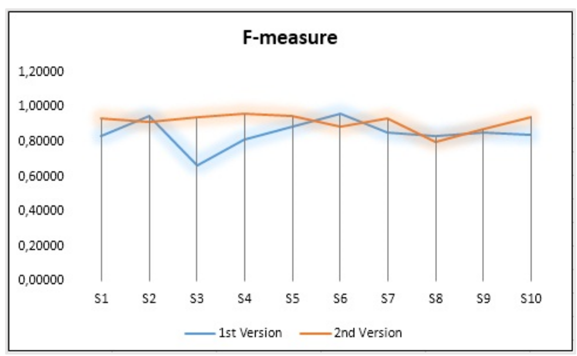

Figure 6: F-measure calculated for all subjects

The **precision** graph describes the profile concepts' relevance to the user, computed by the division of the number of concepts relevant to the user (given by the user) by the profile concepts' number.

The **recall** graph shows how complete a user profile is, calculated by the division of the relevant concepts number in a profile by the relevant concepts number given by the user.

F-measure combines precision and recall, as it is a harmonic mean of both previous metrics. Results confirmed that the user profile's accuracy and relevance rates increase significantly when

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

combining multiple user browsing data sources. However, when the profile concepts' number is important, the rate of irrelevant concepts is perceptible.

Generally, the classifier is proven accurate enough, and most classified concepts are indeed relevant. Also, the noise (irrelevant concepts) caused by misclassification or irrelevant data follows no particular pattern, this can be investigated in the future works.

5.3 Comparative Analysis

The method we propose in this work focuses on creating ontology-based user profiles implicitly from multiple user browsing data sources and also on optimizing the generated user profile by taking various measures including the exploitation of users' feedback.

In this section, we outline the main contributions of this paper compared to previous studies deemed interesting in this field and discussed in section 2.

Though the profiling method we propose shares the subsequent common aspects of previous profiling techniques;

- Implicitly tracks user browsing data as in [9]
- Classifies the collected user browsing data using the VSM (Vector Space Model) that is also used in [12], [20]
- Uses hierarchically-arranged collections of ontology concepts to represent the user profiles as in [25]
- Requires the user feedback as in [19], [26], [25]

It is distinguished from the previous works in various points, which are considered the main contributions of the present paper;

- To obtain an accurate and complete user profile, we combine multiples browsing data sources (Visited web pages, Bookmarks, Search queries, and Click-through data). Authors in reference [12] combined the visited pages and bookmarks ignoring the value of search queries and click-through data as it represents what the user is really looking for. The others only used one browsing data source (the visited web pages mostly).
- Building accurate and complete user profiles is a continuous process and in this work we propose a profile optimization mechanism. Starting from profile stability, concepts ordering/ranking, to exploiting the user's feedback in enhancing his profile. This process helps in maintaining an accurate and dynamic profile that represents the user's reel interests and preferences.
- Most studies discussed in section 2 propose a domain-specific user profiling method. While our method, with the use of a global domain ontology like ODP [1], can be adapted to a variety of personalization applications; such as in E-commerce, E-learning, Personalized search engines, etc.
- Furthermore, we propose an optimized profile structure; by classifying concepts in the user profile to short-term and long-term interests. This separation is based on the concept weight and time; For instance, if a concept is detected in many browsing sessions, over time it's weight increases and the concept is then classified as a long-term interest. Authors in reference [9] also separated profile concepts to short-term and long-term interests, but the separation was based only on the concept's position in the ontology. In more detail, if the concept is parent node then it is considered a long-term interest, while leaf nodes are deemed short-term. This separation is inaccurate since it ignores the concept's importance (weight) to the user and the concept's visiting frequency. For example, if a user visits a football web page according to [9], the concept "Football" is classified as a short-term interest and "Sport" as a long-term one, even if this user is not a "Sport" fan.
- User feedback is highly important, while authors in previous studies ([19], [26], and [25]) require user feedback, generally through surveys, to detect irrelevant interests. We offer users a GUI on which they have full control over their collected profiles as described in section 3.2.4. This choice not only allows us to easily acquire the user's feedback but also to build user's trust in the personalization system and reduce his fair of the privacy problems.

The obtained results from initial experiments prove that our method helps to improve the profile accuracy and increase users' satisfaction. Yet, we intend to evaluate its performance on a larger scale.

6. CONCLUSION AND FUTURE WORK

The work presented in this paper is focused on creating accurate and dynamic user profiles by collecting user's browsed data over time and classifying vectors of the browsed Web pages, bookmarks, and search queries to concepts in a

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

reference ontology (ODP) using the vector space model.

Some optimizations can still be addressed in the future; Irrelevant concepts, for instance, must be efficiently discovered and removed. We can also exploit other user data sources; like social networks to enrich the user profile. In future studies, expanding the subjects in the experimentation phase and investigating whether there are age or gender differences is necessary.

Using the collected user profiles to optimize search results is in our plans [32]. We intend to provide an enhanced personalized search experience on a question-answering system we developed in [30], to deliver personalized answers following the user's interests.

Protecting user privacy is also a major issue in the PWS field, and we plan on addressing this problem in future work employing the Homomorphic encryption, for its promising results on previous works [31][33]. If a PWS is unable to ensure privacy protection, eventually it loses the users' trust, and will only be utilized by a few people to whom personalization matters more than privacy.

REFERENCES:

[1] AOLInc. Open directory project (ODP)., 1998. Accessed: 2019-04-08.

[2] R. A. Baeza-Yates and B. Ribeiro-Neto. Modern Information Retrieval. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 1999.

[3] V. Challam, S. Gauch, and A. Chandramouli. Contextual search using ontology-based user profiles. In Large Scale Semantic Access to Content (Text, Image, Video, and Sound), RIAO '07, pages 612–617, Paris, France, France, 2007. LE CENTRE DE HAUTES ETUDES INTERNATIONALES D'INFORMATIQUE DOCUMENTAIRE.

[4] C. C. Chen, M. C. Chen, and Y. Sun. Pva: A self-adaptive personal view agent. J. Intell. Inf. Syst., 18(2-3):173–194, Mar. 2002.

[5] Chuhan Wu, Fangzhao Wu, Mingxiao An, Jianqiang Huang, Yongfeng Huang, and Xing Xie. 2019. NPA: Neural News Recommendation with Personalized Attention. In Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (KDD '19). ACM, New York, NY, USA, 2576-2584. DOI: https://doi.org/10.1145/3292500.3330665.

[6] K. Swain and A. K. Nayak, "A Review on Rule-Based and Hybrid Stemming Techniques," 2018 2nd International Conference on Data Science and Business Analytics (ICDSBA), Changsha, 2018, pp. 25-29. doi: 10.1109/ICDSBA.2018.00012.

[7] Gina George, Anisha M. Lal. (2019) Review of ontology-based recommender systems in e-learning, Computers & Education, Volume 142, 2019, DOI: 10.1016/j.compedu.2019.103642.

[8] A. Hawalah and M. Fasli. Dynamic user profiles for web personalisation. Expert Systems with Applications, 42(5):2547 – 2569, 2015.

[9] H. R. Kim and P. K. Chan. Learning implicit user interest hierarchy for context in personalization. In Proceedings of the 8th International Conference on Intelligent User Interfaces, IUI '03, pages 101–108, New York, NY, USA, 2003. ACM.

[10] Dimitrios Bountouridis, Jaron Harambam, Mykola Makhortykh, Mónica Marrero, Nava Tintarev, and Claudia Hauff. 2019. SIREN: A Simulation Framework for Understanding the Effects of Recommender Systems in Online News Environments. In Proceedings of the Conference on Fairness, Accountability, and Transparency (FAT* '19). ACM, New York, NY, USA, 150-159. DOI: https://doi.org/10.1145/3287560.3287583.

[11] Brent Smith,Greg Linden. "Two Decades of Recommender Systems at Amazon.com". IEEE INTERNET COMPUTING. 2017. vol 21, pp:12-18.

[12] J. P. McGowan, N. Kushmerick, and B. Smyth. Who do you want to be today? Web personae for personalised information access. In Proceedings of the Second International Conference on Adaptive Hypermedia and Adaptive Web-Based Systems, AH '02, pages 514–517, London, UK, UK, 2002. Springer-Verlag.

[13] S. E. Middleton, H. Alani, and D. D. Roure. Exploiting synergy between ontologies and recommender systems. CoRR, cs.LG/0204012, 2002.

[14] Mohseni M., Maher M.L., Grace K., Najjar N., Abbas F., Eltayeby O. (2019) Pique: Recommending a Personalized Sequence of Research Papers to Engage Student Curiosity. In: Isotani S., Millán E., Ogan A., Hastings P.,

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

McLaren B., Luckin R. (eds) Artificial Intelligence in Education. AIED 2019. Lecture Notes in Computer Science, vol 11626. Springer, Cham.

[15] A. Mitchell. Meetimer, 2006. Accessed: 2019-06-08.

[16] Wang PK., Hong CF., Lin MH. (2019) Interactive Genetic Algorithm Joining Recommender System. In: Nguyen N., Gaol F., Hong TP., Trawiński B. (eds) Intelligent Information and Database Systems. ACIIDS 2019. Lecture Notes in Computer Science, vol 11432. Springer, Cham.

[17] Asnat Greenstein-Messica, Lior Rokach, Personal Price Aware Multi-Seller Recommender System: Evidence from eBay, Knowledge-Based Systems (2018), doi:10.1016/j.knosys.2018.02.026.

[18] B. Mobasher, H. Dai, T. Luo, and M. Nakagawa. Discovery and evaluation of aggregate usage profiles for web personalization. Data Mining and Knowledge Discovery, 6(1):61–82, 2002.

[19] M. Pazzani, J. Muramatsu, and D. Billsus. Syskill & webert: Identifying interesting web sites. In Proceedings of the Thirteenth National Conference on Artificial Intelligence- Volume 1, AAAI'96, pages 54–61. AAAI Press, 1996.

[20] A. Pretschner and S. Gauch. Ontology based personalized search. In Tools with Artificial Intelligence, 1999. Proceedings. 11th IEEE International Conference on, pages 391–398.

[21] A. S. Ramkumar and B. Poorna. Ontology based semantic search: An introduction and a survey of current approaches. In Intelligent Computing Applications (ICICA), 2014 International Conference on, pages 372–376, March 2014.

[22] E. Rich. Readings in intelligent user interfaces. chapter User Modeling via Stereotypes, pages 329–342. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1998.

[23] Puxuan Yu, Wasi Uddin Ahmad, and Hongning Wang. 2018. Hide-n-Seek: An Intent-aware Privacy Protection Plugin for Personalized Web Search. In The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval (SIGIR '18). ACM, New York, NY, USA, 1333-1336. DOI: https://doi.org/10.1145/3209978.3210180.

[24] L. Shou, H. Bai, K. Chen, and G. Chen. Supporting privacy protection in personalized web search. IEEE Transactions on Knowledge and Data Engineering, 26:453–467, Feb 2014.

[25] F. Tanudjaja and L. Mui. Persona: A contextualized and personalized web search. In Proceedings of the 35th Annual Hawaii International Conference on System Sciences (HICSS'02)-Volume 3 - Volume 3, HICSS '02, pages 67–, Washington, DC, USA, 2002. IEEE Computer Society.

[26] C. G. Thomas and G. Fischer. Using agents to personalize the web. In Proceedings of the 2Nd International Conference on Intelligent User Interfaces, IUI '97, pages 53–60, New York, NY, USA, 1997. ACM

[27] M. He, X. Wu, J. Zhang and R. Dong, "UP-TreeRec: Building dynamic user profiles tree for news recommendation," in China Communications, vol. 16, no. 4, pp. 219-233, April 2019.

[28] J. Chen, "The Implementation of a Personalized Reading System," 2018 14th International Conference on Semantics, Knowledge and Grids (SKG), Guangzhou, China, 2018, pp. 235-240. DOI: 10.1109/SKG.2018.00048

[29] Allison J.B. Chaney, David M. Blei, and Tina Eliassi-Rad. 2015. A Probabilistic Model for Using Social Networks in Personalized Item Recommendation. In Proceedings of the 9th ACM Conference on Recommender Systems (RecSys '15). ACM, New York, USA, 43-50. https://doi.org/10.1145/2792838.280019

[30] E El-Ansari, A., Beni-Hssane, A., & Saadi, M. (2017). A multiple ontologies based system for answering natural language questions. In Europe and MENA Cooperation Advances in Information and Communication Technologies (pp. 177-186). Springer, Cham. DOI: 10.1007/978-3-319-46568-5_18

[31] El Makkaoui, K., Beni-Hssane, A., Ezzati, A., & El-Ansari, A. (2017). Fast cloud-RSA scheme for promoting data confidentiality in the cloud computing. Procedia computer science, 113, 33-40. DOI: 10.1016/j.procs.2017.08.282.

[32] El-Ansari, A., Beni-Hssane, A., & Saadi, M. (2020). An improved modeling method for profile-based personalized search. In Proceedings of the 3rd International Conference on Networking, Information Systems &

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Security (pp. 1-6). DOI: 10.1145/3386723.3387874

[33] El-Ansari, A., Beni-Hssane, A., & Saadi, M. (2020). An Enhanced Privacy protection scheme for Profile-based personalized search. International Journal of Advanced Trends in Computer Science and Engineering. Volume 9, No.3, May - June 2020 DOI: 10.30534/ijatcse/2020/241932020