Leveraging Semantic Web Technologies for More Relevant E-tourism Behavioral Retargeting

Chun LU
Sépage & STIH
chun@sepage.fr

Milan Stankovic
Sépage & STIH
milstan@sepage.fr

Philippe Laublet
STIH, Université Paris-Sorbonne
philippe.laublet@paris-sorbonne.fr

ABSTRACT
The e-tourism is today an important field of the e-commerce. One specificity of this field is that consumers spend much time comparing many options on multiple websites before purchasing. It's easy for consumers to forget the viewed offers or websites. The Behavioral Retargeting (BR) is a widely used technique for online advertising. It leverages consumers' actions on advertisers' websites and displays relevant ads on publishers' websites. In this paper, we're interested in the relevance of the displayed ads in the e-tourism field. We present MERLOT 1, a Semantic-based travel destination recommender system that can be deployed to improve the relevance of BR in the e-tourism field. We conducted a preliminary experiment with the real data of a French travel agency. The results of 33 participants showed very promising results with regards to the baseline according to all used metrics. By this paper, we wish to provide a novel viewpoint to address the BR relevance problem, different from the dominating machine learning approaches.

Categories and Subject Descriptors
H.4 [**Information Systems Applications**]: Miscellaneous

General Terms
Algorithms, Experimentation

Keywords
e-tourism; Behavioral Retargeting; travel destination; recommender systems; Semantic Web

1. INTRODUCTION
The e-tourism is today an important field of the e-commerce. According to Google 2013 Traveler [11], more than 80% people do travel planning online. One specificity of this field is that consumers (more than 60%) spend much time comparing many options on multiple websites before purchasing because finding value is important. In average, 45 days are spent and 38 visits to travel sites are conducted before booking [5].

So when people leave a travel website, it doesn't necessarily mean that they aren't interested or don't like the offers of the website. It might just mean that they want to compare with other options. In this stage of travel shopping, it's easy for people to forget the offers or the name of the travel website. These people wouldn't return and they are thus lost. Behavioral Retargeting (BR) is a widely used technique to address this problem. BR is a form of online targeted advertising. It leverages consumers' actions on advertisers' websites and displays relevant ads on publishers' websites.

WWW'15 Companion, May 18–22, 2015, Florence, Italy.
ACM 978-1-4503-3473-0/15/05.
http://dx.doi.org/10.1145/2740908.2742001

We're interested in the relevance of the displayed ads in the e-tourism field. 68% of people begin planning travel online without having a clear travel destination in mind [11]. Our research hypotheses are that the travel destinations have a big impact on the relevance of the displayed ads and by improving the relevance of the travel destinations we can improve the relevance of displayed ads.

The main contribution from this paper is a Semantic-based travel destination recommender system that can be deployed to improve the relevance of BR in the e-tourism field.

The remainder of the paper is organized as follows. In Section 2, we present the background of our work; in Section 3, we present the MERLOT 1 system; in Section 4, we present the conducted experiments; in Section 5, we conclude the paper.

2. Background
BR systems often consist of two main components. The first component is a bidding system that decides whether to display, where to display and for how much. The second component is a recommender system that decides which ads to display.

Much work has been done in the scope of the first component. In [6], the authors considered the problem of estimating user's propensity to click on an ad or make a purchase. They predicted whether a user in a particular session is a *clicker* or just a *browser*.

In [3], a semantic approach is combined with a syntactic one to improve the relevance of ads for the *Contextual advertising*. The authors proposed a novel way of matching advertisements to web pages that rely on a semantic topical match as a major component of the relevance score. The semantic match relies on the classification of pages and ads into a 6000 nodes commercial advertising taxonomy to determine their topical distance. As the classification relies on the full content of the page, it is more robust than individual page phrases. The evaluation demonstrated a significant effect of the semantic score component. The relevance considered in this paper is the relevance of an ad with regards to a web page. The relevance in our work is the perceived relevance of an ad itself.

The second component is less discussed in BR-related papers and is more developed in papers related to recommender systems. The internal functions for recommender systems are characterized by the filtering algorithm. The most widely used classification divides the filtering algorithms into: (a) collaborative filtering, (b) demo-graphic filtering, (c) content-based filtering and (d) hybrid filtering [2]. Criteo [1] is a popular performance advertising technology company whose global reach is placed as second only to Google's Display Network [8]. We didn't find published papers that explain in detail their approach. By consulting their official website and several journalistic articles [1,4,10,12], we believe

[1] http://www.criteo.com/

that they use mainly some machine learning collaborative filtering approach.

[7] evaluates whether indeed firms benefit from targeting consumers with information that is highly specific to their prior interest. The results showed that consumers who have developed narrowly constructed preferences have a greater focus on specific and detailed product information and therefore are more likely to respond positively to ads displaying specific products. [13] uses 7 days' ads click-through log data coming from a commercial search engine to compare different Behavioral Targeting strategies and validate the effectiveness. The experiment results shows that Behavioral Targeting can do a great help for online advertising and using short term user behaviors to represent users is more effective than using long term user behaviors for BR. In this paper, we work in this direction. In the e-tourism case, with consumers frequently navigating, clicking, consulting and comparing, a rich short-term preferences profile is constructed implicitly. Based on the constructed profile, our system tries to display ads having relevant travel destinations.

The semantic approach that we propose calculates the relevance of two informational resources (e.g., persons, books, movies, keywords) by exploiting the paths between those two resources in a semantic graph. Semantic graphs, such as those resulting from publishing of Linked Data are data structures where informational resources of different types, each having a unique identifier – URI, are interconnected with links of different types. In the example given on the Figure 1, we can see the resource Paris connected with the resource France with the link of type "country". This resource is also described with literal values for some of its properties (e.g. population, latitude). Resources sharing the same value for the same property may be considered implicitly linked.

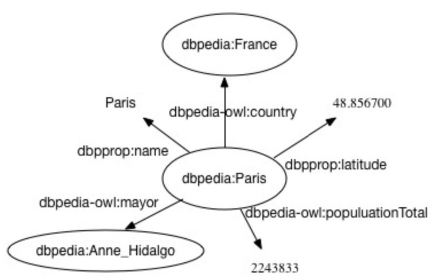

Figure 1. Example part of a semantic graph

3. MERLOT 1 system

In this section, we present the design of the MERLOT 1 system.

3.1 Motivations and Main assumptions

The realization of MERLOT 1 is motivated both by user needs and by the developments in the technological context that offer great promises for the conception of next generation relevance engines. MERLOT 1 should reply to the following user needs:

- Provide plausible alternative destinations to a user's current destination choice. This is especially useful if for some reason the user's destination choice, in the given time interval does not allow the purchase of reasonably priced tickets and hotel arrangements, which often causes frustration with users having trouble to make a proper choice that fits their budget. The system should provide several, diverse enough recommendations in order to offer a relevant choice to the user.

- Leverage publicly available, Linked Data knowledge bases, relevant to travel, to augment the user behavior data, and glean deeper understanding of user actions, thus being able to provide more relevant ads with higher likelihood of click and conversion.

In the context of the exponential growth of public structured data sources, under the initiative called Linking Open Data[2] and with the help of various governmental Open Data initiatives, the Web is now rich with freely exploitable semantic data sources. A considerable number of those sources concern geographical or travel data, usable for calculating travel destination relevance and semantic proximity.

One of the objectives of MERLOT 1 is to leverage the semantic data graphs to generate useful destination recommendations. We assume that the informational richness of those semantic data graphs can help us construct a flexible and versatile approach to destination recommendation, easily adaptable to users' preferences and to the specifics of different scenarios in which the recommendations may take place. MERLOT 1 is not a front-end system. It works in the background of an e-tourism website in order to improve its capacity to better serve the user.

3.2 The relevance calculation & Destination Suggestion Process

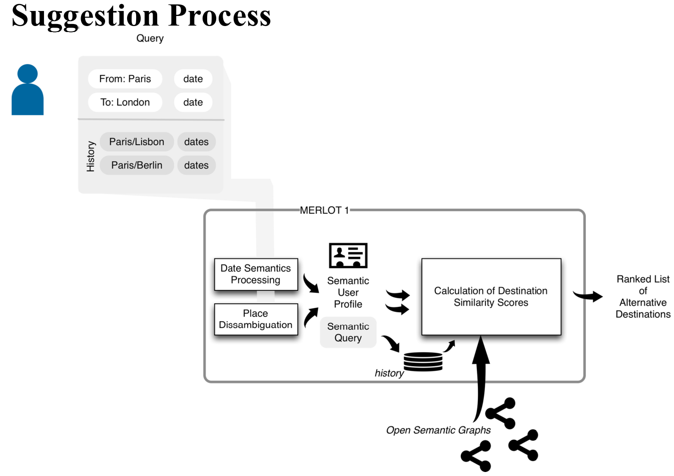

Figure 2. MERLOT 1 system workflow

MERLOT 1 replies to a query submitted by another system (such as a travel website). The query may contain information about the user's context (a destination he is considering to travel to), travel pages that the user viewed and the history of user's previous travel. This input is first treated by MERLOT 1 to transform it to semantic form. Two transformations, detailed in further text, are applied: extraction of date semantics and place disambiguation. The objective of date semantics module is to transform the dates of users desired travel, as well as past travels, to semantic classes such as "weekend", "bank holiday", "summer holiday". Such classes will help adapt the recommendations to the type of trip that the user is interested in. Place disambiguation concerns the transformation of place names, as provided by the user or by the travel website to standardized names used in semantic data sources. In addition to semantically processing the request data, the system stores the received queries as "history" for further use when calculating the semantic proximities. Such a history is by definition more diverse than the history data of any particular website as it contains queries collected from various websites – clients of MERLOT 1. Using historical data, as well as the available semantic data sources, the system then calculates the

[2] http://linkeddata.org/

similarity scores of candidate destinations with the destinations in the user query. We call the similarity measure used by MERLOT 1 the geo-semantic proximity as it calculates both the proximity of two destinations with regards to their geographical distance and with regards to their distance in the semantic graph. We present this measure in more details in further text, as well as the alternative measures that can be used in its place. In the final step the system ranks the alternative destinations with regards to their geo-semantic proximity scores.

In the following subsections we describe the different parts of the recommendation process in more details.

3.3 Input

The query that is provided as an input to the system consists of several trips: a desired trip, for which the user is considering buying a ticket and generally a list of previous trips. Each trip consists of two places – place of origin and a destination, as well as of two dates – the date of ongoing trip and of return trip.

A query does not always have all those elements. For instance, it is possible to imagine a query without the history of past trips in the case of a user new to the system. It is also possible to imagine a query only with the history of past trips in case the user did not yet specify his desired destination, and the travel website of his choice is using MERLOT 1 to generate recommendations in anticipation of his search.

In the context of BR, the input consists of a list of travel offers found in pages previously visited by the user. Existing BR systems already track and rely on such a list of pages that is put in correspondence with a particular e-commerce catalogue of products in order to determine products that the user consulted by browsing the pages from the history.

3.4 Place Disambiguation

The place disambiguation concerns a mapping of keywords (or taxonomic entities) referring to places in the user query to a finite set of concepts used to describe those same places in the semantic data graphs.

$$disambiguate : K \cup T \rightarrow C \quad (1)$$

Disambiguate is a function that provides, for a given keyword (from the set of all keywords K) or a given taxonomical entity from the taxonomy of places T used by a particular travel website, a corresponding concept defined in a semantic graph of concepts C. For instance, for a given keyword "Paris, France", this function would return the concept http://dbpedia.org/resource/Paris provided that DBPedia.org is a chosen semantic graph. This transformation allows to use a unified set of destination identifiers in all calculations and avoid ambiguity and identity problems.

Existing methods, such as concept extractors for text (e.g. Zemanta[3], OpenCalais[4], DBPedia Spotlight[5]), can be used to perform this transformation, and for this reason, we assume that the place disambiguation task is feasible (either by using one of those services or their small adaptations).

3.5 Date Semantics

The phase of date semantics processing consists of determining if any significant date classes appear within the travel period cornered by a trip. For instance if a trip is short and contains a week-end it is indicative that it is a weekend trip, the nature of which is often different than the trips taken in summer for summer holidays. Users often choose different destinations for different types of trips and the motivation behind the discovery of the class of dates is to better understand the likely user intention and adapt the recommendations to it. For the moment we are using the following list of date classes, that we may extend or refine in the future: weekend, bank holiday and summer holiday.

If several classes may be attributed to one time interval we chose only one – the dominant one. We consider that bank holiday is dominant over weekend, and that summer holiday is dominant over bank holiday. The intuition behind this approach is that the less frequent date class is more likely to be the reason for traveling on the given dates than the more frequent one.

Thanks to the available open data sources it is possible to determine the date class automatically. By looking up dates on DBPedia.org it is possible to decide if a particular time interval contains a weekend, a bank holiday in the country of origin. Additionally, longer holidays on places with a beach in summer may easily be labeled as summer holidays. The date class determination is thus a feasible and rather simple task that will prove very useful later on.

In the case the travel offers that the user consulted contained no particular dates, then the data semantics analysis cannot be performed. This step is thus optional.

3.6 Destination Similarity Calculation

Following the Place Disambiguation and Date Semantics Processing phases, the data is transformed in the form of trips (desired and past) in the following form:

$$T = <Po,Pd,Do,Dd,DC> \quad (2)$$

A trip T consists of:

Po - a URI identifying a place of origin in a given semantic graph

Pd – a URI identifying a destination place in a given semantic graph

Do – the date of outward trip (optional)

Dd – the date of return trip (optional)

DC – the dominant date class (optional)

The data consist of two main elements: the semantic user profile, which consists of past trips in the semantic form and the current travel query, consisting of one trip in the semantic form.

The main task of the Destination Similarity Calculation is to calculate the similarity scores of available destination alternatives and rank them with regards to the likelihood that they will interest the user either as alternative destinations to his current travel project, or as inspiring destinations that might motivate him to consider undertaking a new trip.

$$sim(PD,Pc) = \sum_{Pd \in PD} sim(Pd,Pc) \quad (3)$$

The similarity function calculates the similarity of a candidate place destination, Pc, to a set of given place destinations PD. The set PD may contain one destination Pd that the user specified in his query, or may contain several destinations found in the history of user trips. All of them may be taken into account when searching for the most appropriate alternative destination. Ideally, the similarity scores can be precalculated to a certain extent to

[3] http://developer.zemanta.com/

[4] http://www.opencalais.com/

[5] http://dbpedia.org/spotlight/usersmanual

accelerate the operation of the destination recommendation engine.

3.6.1 Hybrid Geo-semantic Proximity Calculation

Hybrid Geo-semantic Proximity (HGSP) is the core destination measure used by MERLOT 1. This measure combines two proximity scores: the score of proximity in the semantic graph of concepts (semprox) and the score of proximity by geographical distance (dist).

$$simHGSP\ (Pd,Pc) = semprox(Pd,Pc) \bullet dist(Pd,Pc) \quad (4)$$

The geographical distance (dist) assures that the suggested destination alternatives are not in disproportion to the distance that the user is willing to travel. Suggesting a traveler from Paris interested in going to Cannes to travel to Hawaii instead would be inappropriate, despite the level of semantic similarity between Cannes and Hawaii destinations. The role of *dist* function is thus primarily in filtering. It augments the score of semantic proximity for places on the similar distance to the initial distance that the user was willing to pass, tolerates places found on a much shorter distance and penalizes the places found on much greater distance. The motivation for such a function is the intuition that users might more likely prefer to travel within similar or shorter distance to those initially planned.

$$dist(Pd,Pc) = \begin{cases} 2, & |geod(Pd,Po) - geod(Po,Pc)| <= \delta \bullet geod(Pd,Po) \\ 1, & geod(Po,Pc) < (1-\delta) \bullet geod(Pd,Po) \\ e^{\left(\frac{1}{|geod(Pd,Po) - geod(Po,Pc)| - \delta \bullet geod(Pd,Po)}\right)} - 1, & geod(Po,Pc) > (1+\delta) \bullet geod(Pd,Po) \end{cases} \quad ,\delta \in [0,1] \quad (5)$$

As shown in equation 5, the *dist* function associates the value 2 to the places Pc found on a similar distance from the place of origin Pc as the place where users initially wanted to travel Pd. This geographical distance is expressed by the function *geod*. A threshold δ is used to define how great the difference in distance can be tolerated for a candidate place to be considered close enough. In our experiments we use δ =0.2, so the places found on a distance 20% shorter or greater than the initial distance (*geod*(Pd,Po)) that the user was prepared to pass, can be considered as close enough. For the places found in a distance much shorter than those specified by the threshold, the dist function asserts the value 1, and considers them as the second best choice. For the places found in a distance greater than the threshold, the dist function attributes an exponentially decreasing value in function of the actual distance. Such places will obtain values of *dist* lower then 1.

The *semprox* fuction quantifies the strength of connections between the initial destination place Pd and the place candidate Pc in a semantic graph. The semantic graph G is composed of a set of informational resources, a subset of which (D) designate places – destinations that are linked with typed links, where T is the set of link types. We define a rather simple graph proximity function, that only takes into account the number and the length of paths that two resources in the graph. To improve the performance of the proximity function in the future it is possible to include more advanced weighting functions. Our function, formalized by the formula 6, calculates the graph proximity of places Pd and Pc in a semantic graph G. We used DBpedia as the semantic graph, but any other graph (or their combinations) may be used in the future. According to our formula the *semprox* proximity of two places is calculated by taking into account all the paths in the graph that exist between these two places. For each path a score is calculated based on the length of the path and the importance of the graph pattern that the path fits into.

$$semprox_G(Pd,Pc) = \sum_{p \in paths(Pd,Pc)} \frac{importance(p)}{length(p)} \quad (6)$$

The importance function allows fine-tuning the approach and giving more priority to places connected over paths that are more indicative of their semantic similarity. It is this function that allows leveraging the semantic nature of links and taking into account the different types of links and paths.

Different importance functions may be used. In the actual version of the system, we used a simple importance function defined by the equation 7. In the future, it would be possible to even further refine this function to focus only on particular link types that prove to be the most significant over time.

$$importance(p) = \begin{cases} 1, & \exists t, t \in T, p \text{ confirms to a pattern } t \\ 0 & \text{otherwise} \end{cases} \quad (7)$$

Our importance function asserts the value 1 to a path p if p confirms to a pattern t from the predefined set of preferred patterns T. Out set T consists of the patterns represented graphically on the figure below. This particular selection is the result of observations of data structures in the DBpedia graph. The chosen graph patterns show best performance and link the most similar destinations to one another.

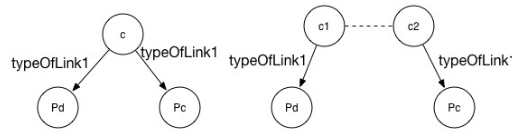

Figure 3. Example of graph patterns

Our patterns cover the paths established between a place Pd and Pc so that, both Pd and Pc are the targets of links of the same type (represent the values of the same property typeOfLink1) reaching from either the same concept (image on the left) or connected concepts (image on the right). In the former case, the concepts can be connected with the paths of variable types and it is the *length(p)* function that will allow us to factor in this length when calculating the score of proximity established over a particular path. An example of an eligible path could be established over a concept that represents an event, for instance Olympic Games, that is connected with the concepts representing Paris and London with the link type "heldIn". While the fact that the same event took place in a particular city some time ago would rarely motivate someone to visit that city, when calculating place similarity this fact may be useful. Similar events are often organized in similar places, and we indeed observed in our data that this pattern is an interesting one to follow.

In practice, we calculate the *semprox* measure by running SPARQL queries on DBpedia to retrieve the candidate places Pc that are findable in the graph proximity of the initial destination Pd, on paths that confirm to our path patterns. Once the place candidates have been collected and the paths lengths calculated it is easy to calculate the *semprox* value.

The calculation of the geographical distance is also possible by relying on DBpedia data about the geographic coordinates of places Pd and Pc. In practice, in order to co-accelerate the calculation process, we perform a single SPARQL query containing both the conditions for *semprox* and *dist* functions.

4. Preliminary Evaluation
Performing a thoroughly complete evaluation would actually require us to implement a bidding system of our own and calculate the click-through rate, which at this early stage of research we are not yet able to do. We thus perform a preliminary user study where we isolate the behavior of our interest and question a panel of Internet users (mixed and aged 23-35) about the likelihood that they click on the ads proposed by our system and by the baseline.

4.1 Baseline
The baseline that we compare with is the popular Criteo engine. Among its clients, Criteo is used by a French travel website, that also uses MERLOT 1 to recommend travel on their website. Both Criteo and MERLOT 1 are installed to track user browsing history on this website, averaging 60k unique visitors a month. The Criteo ads of this advertiser are displayed on the website of "So Foot[6]". MERLOT 1 and Criteo system use two different approaches, this allows us to compare our proposed semantic approach to what we believe is a machine learning collaborative filtering approach, used by the baseline on the same input data and on the same e-commerce offer catalogue.

4.2 Experiment procedure
We established a questionnaire online and asked the participants to perform the following steps:

The main steps that participants followed are:

1. Participants put themselves into the scenario of looking for options for their next travel. They had to imagine a concrete occasion for travel (for instance next holidays).
2. They went to the mentioned French travel website, and were asked to consider at least 3 offers available there (i.e. visit at least 3 pages with travel offers)
3. They were redirected to the So Foot website, where the baseline systems showed its ads on a right sidebar. They evaluated each of the 3 offers according to a Five-Level Relevance Scale.
4. The evaluation form led them to a page similar to So Foot, where on the same placement on the right, 3 offers from MERLOT 1 were placed instead of the baseline ones. They evaluated each of the 3 offers according to the same scale.
5. They gave an overall impression and a free comment.

The system randomly changed the order of steps 3 and 4.

The Five-Level Relevance Scale that we used is:

1. This offer doesn't interest me at all.
2. This offer might be interesting but I don't click.
3. This offer seems interesting but I hesitate to click.
4. This offer is interesting and I click to know more.
5. This offer is very interesting and I click to know more.

4.3 Metrics
We used four metrics to gauge the relevance performance: *Total number of offers corresponding to each level, Precision, Recall* and *F1 score.*

Total number of offers corresponding to each level: In the Relevance Scale, from level 1 to 5, the relevance is in ascending order. The more there are offers corresponding to levels closer to 5, the better relevance performance the system has.

Precision is the number of relevant offers divided by the total number of offers. We consider that offers corresponding to level 1 and 2 as not relevant, offers corresponding to level 3 half-relevant and they count 0.5, offers corresponding to level 4 and 5 are relevant and they count 1 as they correspond to an explicit intention to click.

$$precision = \frac{number\ of\ relevant\ offers}{total\ number\ of\ offers}$$

Recall is modified to adapt to our situation. Since participants evaluated only the displayed ads. We don't know their appreciation about the non-displayed ads. The modified recall is the number of relevant offers divided by the total number of relevant offers of the two systems.

$$modified\ recall = \frac{number\ of\ relevant\ offers}{total\ number\ of\ relevant\ offers\ of\ the\ 2\ systems}$$

F1 score is the harmonic mean of precision and recall.

$$F_1 = 2 * \frac{precision * recall}{precision + recall}$$

4.4 Results & Discussion
33 participants answered our questionnaire. Here are the results of the metrics mentioned above.

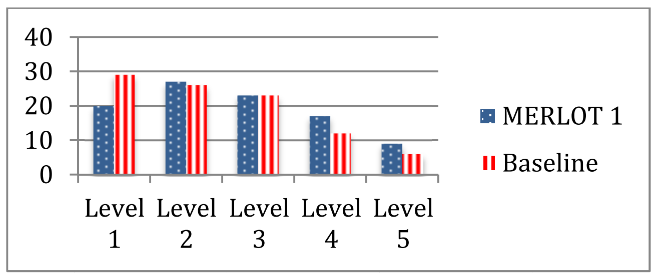

Figure 4. Total number of offers corresponding to each level

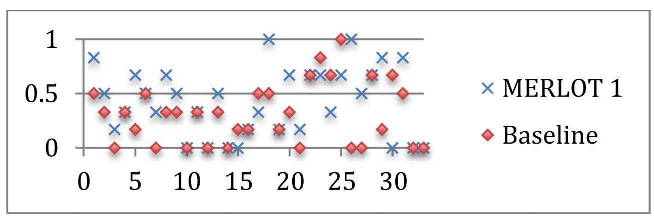

Figure 5. Results of precision metric

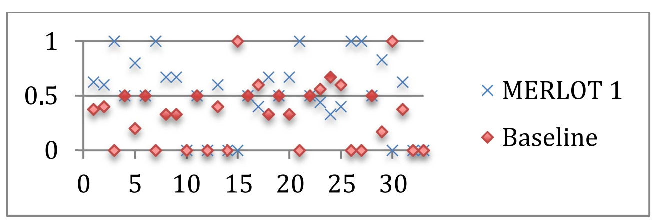

Figure 6. Results of recall metric

[6] http://www.sofoot.com/

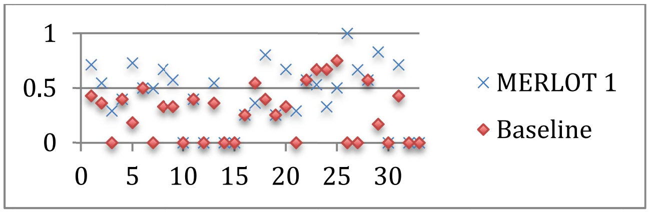

Figure 7. Results of F1 score metric

In Figure 4, we can see a clear advantage for MERLOT 1 in terms of total number of relevant offers (which correspond to Level 4 and 5) – it brings higher chances to generate click-worthy ads. Figure 5, 6, 7 are respectively results of precision, recall and F1 score. Again, MERLOT 1 outperforms the baseline system. The average F1 score of MERLOT 1 offers was 0.431 whereas the baseline had 0.27. Since these results correspond to an average over 3 offers showed to each user at a time in one ad banner (vertically), it is also interesting to observe the difference offer by offer. When looking at the first offer, MERLOT 1 was slightly more relevant with 0.359 (vs. 0.328 for baseline), but it is actually on the second offer where most of the difference is made in favor of our system (0.5 vs. 0.234). This difference can be explained by the fact that the baseline system often proposes an offer already seen by the user (likely to be highly relevant) on the first place, while our system doesn't show the offers already seen. In addition to our main metrics, we were interested to see what is the probability (for both systems) to generate a rating 4 or 5 (one of the ratings corresponding to the intention to click). Such probability was 0.271 for our system vs. 0.177 for the baseline. Obviously these values are too high to be considered an accurate estimate for a probability of click, as it is normal to expect the users to be more generous in clicks in a declarative study then in regular browsing behavior, but we can safely assume that their generosity was equally high for both systems (as same users performed the study, and systems were presented in random order) so it is reasonable to believe that the observed difference may indicate a potential of MERLOT 1 to generate higher click-through rates if used in a real browsing scenario. In the free comment part, some participants found that the offers generated by MERLOT 1 are diverse and surprising, that those generated by the baseline are too similar to the clicked offers and they lack of diversity. The diversity and curiosity that the MERLOT 1's ads show, may explain a part of the differences created in the systems' performance.

5. Conclusions

In this paper, we presented MERLOT 1, a semantic-based travel destination recommender system that can be deployed to improve the relevance of Behavioral Retargeting in the e-tourism field. 33 people participated in the evaluation. MERLOT 1 system outperforms the baseline according to all used metrics. While the time and sample size represent limitations to our study, the convergence of results on multiple metrics, indicates that the use of Semantic Web data to augment behavioral data may be a promising approach to improve the performance of behavioral retargeting systems in the future. After confirming the promising nature of a Semantic Web approach, we will conduct further experiments in a more quantitative setting focusing only on ad click data as opposed to explicit user interrogation that we used in the paper – which has the merit of providing detailed insight, but had to be performed on smaller sample volumes.

6. REFERENCES

[1] "Au cœur du Moteur Criteo". (2015, January). Retrieved from <http://www.viuz.com/2014/04/07/au-coeur-du-moteur-criteo/>

[2] Bobadilla, J., Ortega, F., Hernando, A., & Gutiérrez, A. (2013). Recommender systems survey. Knowledge-Based Systems, 46, 109-132.

[3] Broder, A., Fontoura, M., Josifovski, V., & Riedel, L. (2007, July). A semantic approach to contextual advertising. In Proceedings of the 30th annual international ACM SIGIR conference on Research and development in information retrieval (pp. 559-566). ACM.

[4] "Criteo : les Big Data au cœur du reciblage publicitaire". (2015, January). Retrieve from <http://pro.clubic.com/webmarketing/actualite-538096-marketing-paris-criteo.html>

[5] "Custom Research: Exploring the Traveler's Path to Purchase". (2014, august 28). Retrieved from <http://info.advertising.expedia.com/path-to-purchas>

[6] Djuric, N., Radosavljevic, V., Grbovic, M., & Bhamidipati, N. (2014). Hidden Conditional Random Fields with Distributed User Embeddings for Ad Targeting. In IEEE International Conference on Data Mining (ICDM).

[7] Lambrecht, A., & Tucker, C. (2013). When does retargeting work? information specificity in online advertising. Journal of Marketing Research, 50(5), 561-576.

[8] "Latest comScore Report Reveals that Criteo Ads Reach Virtually One Billion Users Worldwide". (2015, January). Retrieved from <http://www.criteo.com/news-and-events/press-releases/2014/10/latest-comscore-report-reveals-that-criteo-ads-reach-virtually-one-billion-users-worldwide/>

[9] Liu, K., & Tang, L. (2011, October). Large-scale behavioral targeting with a social twist. In Proceedings of the 20th ACM international conference on Information and knowledge management (pp. 1815-1824). ACM.

[10] "The Criteo engine". (2015, January). Retrieved from http://www.criteo.com/what-we-do/technology/

[11] "The 2013 Traveler". (2014, August 28). retrieved from <http://www.thinkwithgoogle.com/research-studies/2013-traveler.html>

[12] "Tout ce qu'il faut savoir sur Criteo… avant son entrée en bourse le 30 octobre". (2015, January). Retrieve from <http://frenchweb.fr/tout-ce-quil-faut-savoir-sur-criteo-avant-son-entree-en-bourse/114500>

[13] Yan, J., Liu, N., Wang, G., Zhang, W., Jiang, Y., & Chen, Z. (2009, April). How much can behavioral targeting help online advertising?. In Proceedings of the 18th international conference on World wide web (pp. 261-270). ACM.