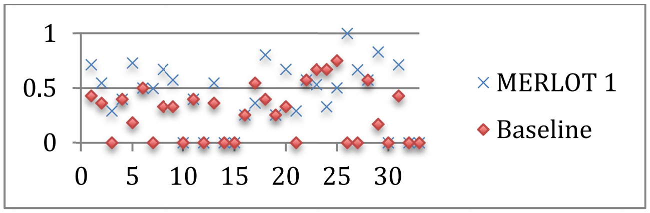

Figure 7. Results of F1 score metric

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.