

A hybrid recommendation system for news in a mobile environment

Paula Viana
Polytechnic of Porto, School of Engineering
and INESC TEC
Porto, Portugal
paula.viana@inesctec.pt

Márcio Soares
INESC TEC
Porto, Portugal
mmsoares@inesctec.pt

ABSTRACT

Over the last few years consumption of news articles has shifted more and more from the written versions towards the web. Mobile devices, which became more powerful, with larger screens and connected to the Internet, have had a great influence on this paradigm change. A critical problem associated to online news is related to the fact that the large number of daily articles can be overwhelming to the users. Recommendation services can largely improve the efficiency and accuracy of acquired information. These systems are designed to filter critical news, key events and meaningful items that might be of interest to a reader.

In this paper, a news recommendation system in a mobility scenario is presented. The implemented recommendation system combines content-based and georeferenced recommendation techniques. Recommendations are supported by short-term and long-term user profiles created implicitly and considering also the mobile device geolocation. The final recommendation list is obtained by combining recommendations provided by the different recommendation approaches. To evaluate the performance of the solution, a user study was conducted. Results indicate that the quality of the recommendations is acknowledged by the test users. The system was integrated in a mobile application of a Portuguese newspaper (Público) in the context of the project Pglobal.

CCS Concepts

• Information systems → Clustering • Information systems → Content ranking • Information systems → Information retrieval • Information systems → Information extraction • Information systems → Geographic information systems • Human-centered computing → User studies

Keywords

Online news; hybrid recommendation system; geolocation; mobile application; user profiles.

1. INTRODUCTION

Online news reading has become very popular as the web provides access to news articles from millions of sources around the world.

People typically read news to know and understand what happened, what is happening and will happen in a town, region, country, or in the world. This information demand presents itself a key challenge for news providers on helping users finding, fast and accurately, those news which are of interest for them. Manually searching and analysing the available news articles, to select those considered interesting, is not feasible within the time constraints common for most users. In this scenario, recommendation systems stand out as a possible solution to help newspaper readers on the selection of content that best fits his/her preferences. Additionally, assisting tools that guide the reader in the content selection process can contribute to increase his loyalty to a specific service/product and thereby contribute to improved cost-effectiveness for the service provider.

The dissemination of mobile devices like smartphones and tablets on everyday activities introduced an additional shift on the reading habits from the online reading paradigm in the computer screen to everywhere, anytime and from any mobile device. Due to the growing use of mobile devices, research on recommendation approaches appropriate for this emergent scenario has been recently addressed by the academic community [7],[27],[41]. However, despite this increasing awareness, the field of mobile news articles recommendations is still fairly unexplored.

Mobile and ubiquitous computing provide extended means to sense data and enhance recommendations that can deliver interest topics to the user, in any place, at any time, for example, while traveling on a business trip or on holidays. As they can provide information on the geographic location of the reader, this information can also be used by recommendation systems to provide news articles with the same geolocation.

In this paper, a hybrid recommendation system for news in a mobile environment is presented. The implemented system was integrated in a mobile application of the Portuguese newspaper Público in the context of the project Pglobal. Pglobal is a new concept of editorial content aggregation that joins in the same platform a network of newspapers, universities and users from various Portuguese-speaking countries. The goal is to build a community of readers who have in common the language and to create or consolidate a payroll for collecting digital content and subscriptions and explore the information business in the international Portuguese-speaking market.

The implemented hybrid recommendation system combines content-based and georeferenced recommendation methodologies. Using these approaches, recommendations are generated based on a user's profile created implicitly by monitoring the user interactions with the mobile application, in particular the news that he reads, shares on social networks or that he Likes. Additionally, the geolocation of the mobile device is also implicitly registered.

Since the reader's interests may change over time, the process of observing his reading habits to enable adapting the output accordingly is performed by maintaining a short-term and a long-term profile. Using the user's two profiles, the system is able to recommend two types of items: (1) items similar with the ones with which the user has recently interacted with, associated to the short-term profile; (2) items containing the terms/tags most read by the user, associated to the long-term profile. Additionally, the geolocation of the mobile device is also used in order to recommend news articles related to the same location. The final recommendation list is obtained by combining recommendations provided by each of the approaches.

The solution was evaluated by a group of volunteers and results indicate that the quality of the recommendations is acknowledged by those test users.

The rest of this paper is structured as follows. In Section 2, we survey the available literature on various state-of-the-art news recommender system approaches. Section 3 gives an overview of the overall system architecture and component interaction. Section 4 describes the news pre-processing stages in order to extract key information from the content. Section 5 focuses on the hybrid recommendation approach, detailing individually each of the methodologies. In Section 6 the proposed solution is evaluated by performing a user study. The results are presented and discussed in detail in Section 7. Finally, in Section 8 some conclusions are taken and future work is highlighted.

2. RELATED WORK

Recommender systems are systems with the ability of providing suggestions or directing a person to a service, product or content, that has a potential of interest among a number of different alternatives. Currently, highly rated internet sites as Amazon¹, eBay² or YouTube³ use recommender systems as part of their services. Beyond these examples, recommender systems also apply on numerous other domains including books [8], movies and TV programs [35],[36], music [17],[18], news articles [4] and so forth.

Over the years, recommendation of online news articles has become an area of great interest. For instance, large newsfeed portals, such as Google News, and Yahoo! News, provide personalized news recommendation services for a large amount of online users.

The task of recommending news articles based on the user's preferences, can be conducted using distinct methodologies. Approaches adopting content-based and collaborative filtering are widely used by existing news recommender systems.

Some content-based news recommender systems have been proposed in the last decade. An example is NewsDude [2] that presents news stories to the user, who then rates the articles according to whether they are interesting or not. The user profile is then compared with content of other news stories to generate personalized recommendations. Another content-based approach is discussed by Newsjunkie [14]. This solution analyses the dynamic of novelty in recommendations and personalized news stories are generated by identifying the novelty of news' topics. Another recommender, presented in [25], proposes a method utilizing a selection of parameters to estimate news relatedness. These parameters include relevance, novelty, connectivity, and transition smoothness. Recommendations are estimated by finding the news

with greater relatedness to the news articles in the users' profiles. A content-based approach with some new contributions is presented in [3]. In this work the authors propose to enhance effectiveness of news recommender systems by adding, to each recommendation, an explanatory statement to help the user better understanding if, and why, the item can be of interest. Experiments show that news recommender systems can greatly benefit from using explanation modules.

In the last few years, some content-based approaches start focusing on exploring news semantic concepts in order to improve recommendation's quality. A news item often contains key concepts that capture the semantic context of the article. Recommenders that focus on the key concepts might produce faster and more accurate recommendations than the content-based recommenders, since they don't need to consider all words and, unlike words, concepts are not ambiguous. Several examples of semantic-based approaches are discussed in [10], [15],[19],[33],[37].

Collaborative recommendation systems consider that users with similar reading behaviours in the past usually will have similar preferences about news articles in the future. Grouplens [32] was the earlier representative of a collaborative filtering approach in the news domain. In Grouplens, users rate articles after reading them and this information is used to correlate users with similar scores. Afterwards, based on ratings from similar users, the system predicts how well users will like new articles. Google News [9] is another popular example of a news articles recommender based on collaborative filtering. Google News is an online news portal that aggregates news articles from thousands of sources, grouping them to the users, according to their personal interests. The recommendation approach is based on implicit user and community feedback represented by the user click history. A click on an item is interpreted as a positive rating.

Content-based and collaborative filtering can provide meaningful recommendations. Each of the approaches have however some disadvantages [36]. In order to improve the performance, hybrid approaches to news recommendations have also been explored. Representative examples include [6],[12],[28].

P-Tango [6] presents a hybrid approach that recommends news items by combining content-based and collaborative filtering recommenders together using a weighted average function. In [28] an adaptive hybrid recommender system is proposed. The term adaptive hereby refers to the fact that the system supports more than one collaborative recommender algorithm. In [12] the authors present a hybrid personalized news recommender system that recommends interesting news articles to the user using the micro-blogging service *Twitter*.

Although news recommendation has been significantly explored in recent years, most research was conducted considering only user interaction with a computer screen. Approaches directly related to mobile devices have only, over the last few years, started to appear. Representative examples include [22],[34],[39].

MONERS [22] presents a mobile web news recommendation system, which determines users' preferences for news articles based on the importance of the news event, the recency of the

¹ <http://www.amazon.com>

² <http://www.ebay.com>

³ <http://www.youtube.com>

article, changes in users' preferences, user segments and article preferences.

The proposal described in [34] presents ePaper, a research prototype system of a personalized newspaper on a mobile reading device which aggregates content from several different news providers. The ePaper classifies the news items according to concepts from news domain ontology, and delivers an electronic newspaper to each subscriber according to the user's profile.

The additionality that mobile devices bring into recommendation systems is that they enable knowing the user's current location. This can be used to provide e.g. the vicinity of restaurants, entertainment and shopping information, etc. Despite the existence of some georeferenced recommendation systems [16],[26],[40], mobile news recommendations based on geolocation remains an unexplored area.

Work presented in [39] describes one of the few mobile news recommendation systems using this functionality. This proposal describes a personalized mobile news recommender system which combines various methods of recommendation, particularly designed for the domain of news articles. It considers a variety of factors including freshness and users' preferences as the main recommendation targets. According to the authors, the information about location is mainly used to alleviate cold-start problems.

The solution being described in this paper uses the concept of geolocation in three domains: the current location of the reader obtained by the device sensors; the identification of the original source of a news item and respective country; the identification of the place the text refers to through geo-tagged information attached to the content.

3. RECOMMENDATION SYSTEM ARCHITECTURE

Figure 1 depicts the architecture of the developed system. Four main blocks can be identified: *Pglobal News Provider*, *Mobile Device*, *Device logs Server* and *Recommendation System Core*.

Pglobal News Provider is responsible for aggregating and providing, news articles from multiple Portuguese language newspaper sources to the recommendation system and to the mobile device. Periodically, the recommendation system module uploads the latest news articles and processes and stores them for future use.

The *Mobile Device* runs the news App responsible for presenting generic news and recommendations to the user.

The *Device Logs Server* is responsible for storing device logs. The logs information contains data about the user's interactions with the App (the identification of the type of interaction and the date when it was performed). A representative example of possible interactions includes sharing or reading news. The logs information contains also the device ID as well as the code name of the country where the device is located. Periodically, the recommendation system module uploads device logs information to process and create or update devices' profiles.

The *Recommendation Core* is the main component of the system. This module has, as input, previously processed news and devices' profiles. All recommendations are periodically generated in offline mode and the final produced recommendations are sent to the *Pglobal repository* in order to be made available to a mobile device.

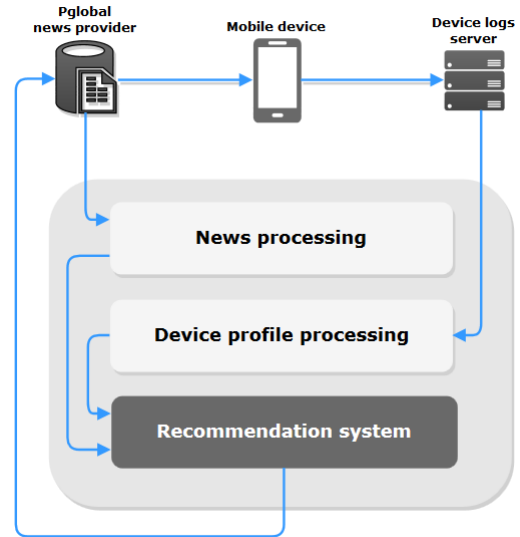


Figure 1. News recommendation system architecture

4. NEWS PROCESSING

Typically, online newspaper articles have a set of metadata to describe them. This may include information such as a title, subtitle, author or date.

The news processing module uses the following metadata elements:

- *title* – title of news;
- *subtitle* – subtitle of news usually introducing key/relevant information;
- *content* – the main text of the news;
- *tags* – tags assigned to a news article that capture key concepts and main topics.

4.1 News pre-processing

In order to extract key information, the system first pre-processes the news article. The first step is to split the title and subtitle into tokens. This is an important prerequisite for various linguistics tasks. In a second step, all common non informative words, defined in literature as stop words, are removed. The next step involves stemming text, reducing inflected words to their *stem* (or root) word. For the main content of the news, despite stemming and removing stop words enables simplifying the text by reducing the number of textual elements, news articles can still be very long. In those situations, the remaining text can still be too long and complex in order to capture the key topics of the text. To retain just the essential information, a filtering process, involving all tags assigned to the news article, is performed. In this step, all words that do not match existing tags are removed.

After this process, the weight of all the remaining words, in all metadata fields, is estimated. This weight shall represent the importance of that term and is estimated combining the number of times that the term occurs, together with the relevance weight of the field where it occurs: words appearing in the title of a news story will have a greater weight compared to a word that can occur many times, but only in the main text of the news. The application of this metric produces a vector where each word in the news article has a weight associated that is estimated according to Equation (1).

$$TermWeight = \left(\frac{TermOccurrences}{TotalTerms} \right) max(TermBoost) \quad (1)$$

where *TermOccurrences* is the number of term occurrences in all news article fields; *TotalTerms* is the total number of terms and *TermBoost* is the metadata field relevance weight where the term occurs. Given that the mobile device has limited space for showing the information, the App will at a first stage show a partial view of the news. The relative importance of each of the fields reflects this situation: Title is attributed an impact of 4, Subtitle of 2, while Content is classified as a 1. If the term occurs simultaneously in the three metadata fields, the *max* weight is considered.

4.2 Clustering

Depending on the news source, dozens of new items can be introduced in the system daily. Estimating the similarity of all these new items with those already stored in the system can result in a very time consuming computational task. Grouping articles according to their content in a set of clusters that contain similar items can contribute to decrease the calculation time of similarity.

The implemented clustering process is supported by the *Carrot2 Framework*⁴. This framework can automatically organize small collections of documents, e.g. search results, into thematic categories. From the list of available approaches the Lingo algorithm [29] was selected given that, according to the framework documentation [30], it is the algorithm that best fits the news domain.

Figure 2 illustrates the result of applying the clustering process to news in the Pglobal systems. Each of these clusters has a set of news that are clearly related to the identified cluster topic.

4.3 Similarity computation

Any content-based recommender system tries to estimate how similar items are. By first defining clusters, we can lower the time needed for similarity estimation. Instead of computing the similarity between all existing items, our system only computes the similarities between items belonging to the same cluster. This methodology avoids calculating the similarity between items that do not have any content in common, increasing the system efficiency. The similarity between items inside each cluster is estimated using the items' weight terms vectors by applying the Cosine similarity measure.

4.4 Device Profile

Users' profiles are the core of many recommendation systems [1],[12],[13]. These profiles may consist in a number of different types of information. For the online newspaper domain, the user profile might contain the type of news or categories the user likes to read. This information can be collected either explicitly or implicitly [35].

A key issue for news recommendation is how to capture each user's reading interests according to his/her historical consumption. Additionally, it is important to take into account the evolution of users' reading interests. According to [24], the topics a user is interested in, can be relatively stable or vary slightly in a long-term perspective, whereas the content accessed by the user might change along different short-term perspectives.

In order to consider the evolution of users' reading preferences, a user short-term and long-term profile was implemented.

In our implementation, the user's profile is constructed by collecting information implicitly, from the observation of the user's actions. All the collected information is stored in a *log* file that is

periodically sent to a web server that processes and stores this information.



Figure 2. Clustering of news

Typically, profile construction and maintenance relies on the history of logged-in users. This might cause potential privacy issues to the users and prevent him from using the system. Our work departs from this restriction and considers only a device profile, which does not require the user login. However, given that mobile devices are usually personal and not-transferable, the collected information is expected to appropriately describe the user.

Implicit information results from three types of user's interaction with the mobile application:

1. *Read* –The read action is registered when a user keeps in the same news screen for a pre-defined period of time. For example, more than 10 seconds.

According to [38] it should be noted that the amount of time that an object is displayed does not necessarily correspond to the amount of time that the object is viewed. However studies like [23] claim that predictions based on time spent reading are nearly as accurate as predictions based on explicit numerical ratings, which is a good indicator for using this approach.

⁴ <http://project.carrot2.org/index.html>

In order to capture implicit feedback about the user's interest in a news article, other approaches can be considered. Representative examples include mouse click [20], scrolling [5] or eye tracking [31]. However each of these approaches has some limitations [10]: using scrolling as an indication of interest may fail, for example, in short texts as scrolling may not be necessary. Additionally, scroll speed should also be considered because, if a user scrolls down quickly on a page without focused reading, the action should instead be interpreted as a sign of lack of interest.

2. *Like* – This action corresponds to a Facebook like that a user gives to a news article.
3. *Share* – This action corresponds to sharing a news article on an available social network, namely, Facebook, Twitter and Google+.

4.5 Long-term profile

The Long-term profile is created based on the news that a user reads over time. For that, the system collects the tags associated to these news and the profile is created by considering the number of times that a specific tag has been read by a user. For example, if a user reads 10 news containing the tag “football” that tag will have a count of 10.

4.6 Short-term profile

The Short-term profile tries to capture recent or seasonal interests, as very popular news. The thought behind this type of profiles is that interest is driven by curiosity or short-term happenings and can quickly become uninteresting. This aspect is also relevant in a mobility scenario, as users will probably modify their reading habits during short stays in other locations.

In this work, the short-term profile is based on the set of recent news with which the user has interacted (read, like or share) with and that were still not considered and processed by the recommendation system.

5. HYBRID RECOMMENDATION APPROACH

Hybrid recommendation systems combine two or more recommendation methodologies.

In this work, the implemented hybrid recommendation algorithm combines content-based and georeferenced recommendations approaches. The final recommendation list is constructed by combining all recommendation lists resulting from each of the recommendations' approaches.

5.1 Content-based recommendations

Content-based recommendation identifies similar items to those that the target user has liked in the past.

In this work, a content-based approach is used to generate recommendations by matching short-term and long-term profiles with news articles information.

For the long-term profile, the algorithm matches the tags that have been cumulatively added to the profiles with the tags of existing news.

For the short-term profile, actions accomplished by the reader (read, like or share) are considered to have different levels of relevance. For example, if a user shares a news story on Facebook, this news is likely to have had more impact on the reader than a

story that he just read without sharing it. Table 1 lists the available actions associated to its relevance weight. The higher the action weight the higher the importance of the action. For each action, a list with the N most similar news is created.

Table 1. Action weights

Action name	Action weight
Read	1
Like	2
Share	3

In order to generate the final recommendation list, all the individual lists are combined. Each list is first sorted according to the action weights - higher action weights having a better position. The final list gets one element from each of the individual lists and the process keeps going until the desired length is achieved.

5.2 Geo referenced recommendations

The Georeferenced approach recommends news having the same geolocation of the mobile device.

Although the geolocation module of the mobile device is able to return very precise information that enables identifying the city or even the street name, our implementation only considers the country where the device is located. The reason for this is that news hardly have very detailed metadata schemas on location. Our approach can be refined and granularity of the location enhanced if news providers start following a strategy of more detailed description/identification of the event.

Given that Pglobal aggregates content from different newspapers in the world, geolocation is considered in two facets: geographic location (country) of the news provider and identification of the location of the event being described in the content.

Geolocation information is coded using ISO 3166-1 Alpha-2 country code.

5.3 Final hybrid recommendation list

The final recommendation list is created by combining all the previous described approaches with a pre-defined fraction for each.

Additionally, a priority is assigned to each of the recommendation methodologies. Whenever one of the recommendation approaches cannot fill the final list in the required percentage, the recommendations will be provided from the following recommendation methodology in the priority list.

The system was tested with selected beta-tester users. Profiles were created and a list of recommendations that reflects all the approaches was created.

Table 2 presents a summary of the information collected from a device that enables creating the user profiles. Information has been processed to reflect the approaches used for short-term and long-term recommendations. Additionally, information on the location is also presented.

For the Short-term profile field, Table 2 presents the most relevant set of stemmed terms extracted from the news with which the user has recently interacted with. The terms in bold identify the ones to which a match was found with other existing news that are then candidate to integrate the final recommendations' list.

The long-term profile section presents the list of the top 20 tags associated with the news the user read in the past.

Table 2. Example of a user profile

Geolocation	br
Short-term	<p>produit, emergenci, alergi, cheg, legislativ, violenci, entregar, cheg, doent, laranj, rajad, amig, deve, forn, natalidad, castel, avis, ficca, lisbo, distrit, opinia, empres, burundi, joao, encenaca, miam, famili, saud, integrad, gesta, farmaci, beja, teatr, port, director, alemanh, integrad, sofisticad, cart, franc, espanh, horari, vesper, mandat, alun, sociedad, maxim, psd, sent, politic, politic</p>
Long-term	<p>PSP(5), Sociedade(5), local(4), Lisboa(4), Câmara do Porto(3), Opinião(3), São João(3), Ministério Público(2), Ministério da Educação(2), Violência(2), Mundo(2), Educação(2), Conflitos(2), PSD(2), Meteorologia(2), França(2), Universidade Lusófona(2), Universidade Lusíada(2), Moldávia(2)</p>

Figure 3 shows an extract of the resulting recommendation list considering the user profile described. In this list, items related to each of the approaches are presented. Words in red are shown to clarify the reason that justifies the news to have been chosen as a recommendation.

6. EXPERIMENTAL EVALUATION

In order to evaluate the recommendation system, a user study was conducted by recruiting a set of volunteer subjects, with the only requirement of being readers of the newspaper, and asking them to perform two distinct tasks.

The user evaluation group consisted of about 35 participants with different reading profiles: (1) seasonal readers that had only read a very small number of news; (2) average readers having an intermediate amount of news accesses; (3) intensive readers, which have a high reading historical profile and are probably the kind of reader that consumes news every day.

Users were asked to perform the following evaluations and to provide their feedback:

1. The first experiment consisted on presenting to the user three lists of news corresponding to (1) the output of the recommendation system, (2) the most popular news and (3) randomly selected news. Each list was composed of 5 newspaper news, horizontally aligned. For this task the user was requested to select the list of news he thinks that best matches his preferences. The ranking position of each list was randomly generated in order to dismiss possible ranking influence. The main purpose of this experiment is to evaluate the impact of the recommendation algorithm on users' satisfaction when compared to the other two simple approaches.
2. For the second evaluation task the user was shown 12 news resulting from the 3 recommendations approaches (Geo-located, Short-term, Long-term) and asked to rate, 1 to 5, each. Each algorithm generated the same number of recommendations. The main purpose of this experiment is evaluating the effectiveness of each of the approaches on the user's satisfaction.



Figure 3. Recommendation list aggregating multiple recommendation approaches

7. RESULTS AND DISCUSSION

Figure 4, 5 and 6 summarize the results obtained in the first experiment.

Figure 4 (resulting from the feedback of the first experiment) confirms that the list resulting from the recommendation algorithm was most often chosen, obtaining almost 50% of the total number of selections. The most popular news list comes afterwards, with 29% of the selections and finally the random list. Only a small difference is observed for these two lists, showing that not always the users enjoy popular news. Results are consistent independently of the position of the recommendation list in the experiment, as depicted in Figure 6, showing that there is no bias on the analysis.

Figure 5 further strengthens the good acceptance of the developed system when considering only loyal readers. Given that volunteers had a large range of reading histories, some of them with a very short record, the analysis presented includes only readers with more than 20 read news.

These results certify the good performance of the developed recommendation algorithm showing that users will more likely read recommended news instead of popular news or other generic content that shows up on the screen. This is expected to contribute to the user's satisfaction and fidelity to the system since he did not have to waste too much time searching for news of his interest.

Figure 7 presents the results obtained by the accomplishment of task 2, indicating that the short-term recommendation algorithm obtained the highest preferences (rating average). These results demonstrate users' higher interest for news similar to those ones he read recently. The low performance of the georeferenced algorithm may be due to the use of only generic information about the country where the user is located. This is a current limitation that might be improved in the future if additional information is introduced by the news provider.

Results also contributed to define the hierarchy of each of the approaches in the list of recommendations presented to the user. Given these results, the short-term algorithm is assigned the first ranking positions with a higher percentage of news. Recommendations provided by long-term algorithm are presented in the intermediate position, while the georeferenced algorithm will provide a smaller number of items in the final collection.

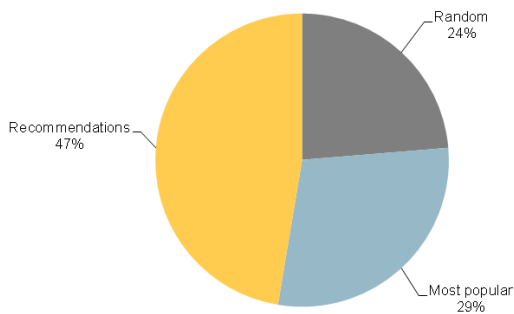


Figure 4. Analysis of the recommendation algorithm when compared to simple approaches

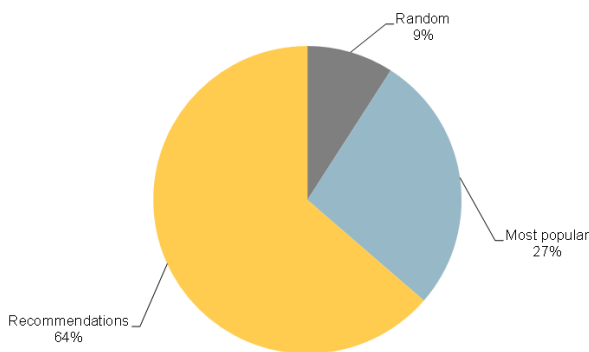


Figure 5. Analysis of the recommendation algorithm when compared to simple approaches for users with a significant record of accesses

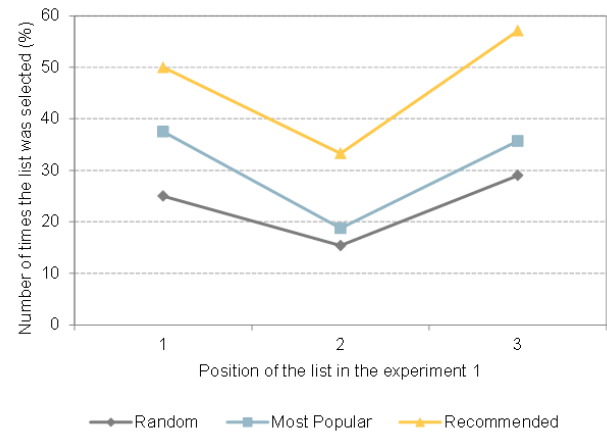


Figure 6. Analysis of the influence of the position of the list in the user selection



Figure 7. Rating average associated with each recommendation algorithm

8. CONCLUSION

The interest on online newspapers has been growing significantly over the past years. In order to present the most relevant news articles to users, different recommendation systems have been made available using various techniques in order to make access to large amounts of information more efficient.

Mobility, social networking and the large number of news providers, bring new challenges but also new opportunities to enhance access to information and improve the user experience when browsing newspapers. This paper presents a hybrid recommendation system for news in a mobile environment that takes into consideration aspects that reflect this new paradigm.

The described approach combines content-based and georeferenced recommendation methodologies. The system is able to collect information implicitly and to create short-term and long-term user profiles that enable recommending: (1) items similar with those the user has recently interacted with (read, share or like); (2) news containing some of the terms/tags most read by the user. Additionally, the location of the mobile device is also implicitly registered in order to recommend news associated to that location.

To the best of our knowledge, this is the first implementation of a news recommendation system for mobile environments that uses the sharing and like of news in social networks to implicitly infer in a more accurate way the user preferences.

The user study acknowledges the good performance of the developed system. Moreover, the results further demonstrated that preferences are mostly influenced by recent interactions than by long last user profiles.

Future work includes evaluating the performance of our implementation using a larger set of users in a real environment. In order to evaluate the level of satisfaction with the recommendations suggested by the system, we will analyse the percentage of recommendations read by the user considering the total amount of news he read. Another experience will involve changing or switch the priority value and the percentage that each of the recommendation approaches provides to the final list and analysing the impact. With these analyses it will also be possible to verify how the user reading behaviour changes and try to infer what is the preferred approach.

ACKNOWLEDGMENTS

The authors would like to thank all the participants in the project Pglobal and in particular Publico for having provided all the grounds and news that enable testing the system in a near-real environment. The work was partially supported by Project NORTE-01-0145-FEDER-000020 financed by the North Portugal Regional Operational Programme (NORTE 2020), under the PORTUGAL 2020 Partnership Agreement, and through the European Regional Development Fund (ERDF).

REFERENCES

- [1] Balabanović, M., Shoham, D. 1997. Fab: content-based, collaborative recommendation, *Communications of the ACM* 40, 3 (March 1997), 66-72. DOI= <http://doi.acm.org/10.1145/245108.245124>
- [2] Billsus, D. Pazzani, M. J. 1999. A personal news agent that talks, learns and explains. In *Proceedings of the third annual conference on Autonomous Agents* (Seattle, Washington, USA, 1999). AGENTS '99. ACM, New York, NY, 268-275. DOI= <http://doi.acm.org/10.1145/301136.301208>
- [3] Blanco, R., Ceccarelli, D., Lucchese, C., Perego, R., Silvestri, S. 2012. You should read this! let me explain you why: explaining news recommendations to users. In *Proceedings of the 21st ACM international conference on Information and knowledge management* (Maui, Hawaii, USA, 2012). CIKM '12. ACM, New York, NY, 1995-1999. DOI= <http://doi.acm.org/10.1145/2396761.2398559>
- [4] Borges, H. L., Lorena, A. C. 2010. A Survey on Recommender Systems for News Data. In *Smart Information and Knowledge Management*, Szczerbicki, E., Nguyen, N. T., Ed. Springer, Heidelberg, 260, 129-151. DOI= [10.1007/978-3-642-04584-4_6](http://doi.acm.org/10.1007/978-3-642-04584-4_6)
- [5] Claypool, M., Le, P., Wased, M., Brown, D. 2001. Implicit interest indicators. In *Proceedings of the 6th international conference on Intelligent user interfaces* (Santa Fe, New Mexico, USA, 2001). IUI '01. ACM, New York, NY, 33-40. DOI= <http://doi.acm.org/10.1145/359784.359836>
- [6] Claypool, M., Gokhale, A., Miranda, T., Murnikov, P., Netes, D., Sartin, M. 1999. Combining Content-Based and Collaborative Filters in an Online Newspaper. In *SIGIR Workshop on Recommender Systems: Algorithms and Evaluation*. (SIGIR '99). Berkley, CA.
- [7] Chen, C.-M., Yang, Y.-C. 2010. An Intelligent Mobile Location-Aware Book Recommendation System with Map-Based Guidance That Enhances Problem-Based Learning in Libraries. In *Advances in Neural Network Research and Applications*, Zeng, Z., Wang, J., Ed. Springer, Heidelberg, 61, 853-860. DOI= [10.1007/978-3-642-12990-2_99](http://doi.acm.org/10.1007/978-3-642-12990-2_99)
- [8] Cui, B., Chen, X. 2009. An Online Book Recommendation System Based on Web Service. In *Proceedings of sixth International Conference on Fuzzy Systems and Knowledge Discovery* (Tianjin, 14-16 August 14 - 16, 2009). IEEE Press, New York, 520-524. DOI= [10.1109/FSKD.2009.328](http://doi.acm.org/10.1109/FSKD.2009.328)
- [9] Das, A. S., Datar, M., Garg, A., Rajaram, S. 2007. Google news personalization: scalable online collaborative filtering. In *Proceedings of the 16th international conference on World Wide Web* (Banff, Alberta, Canada, 2007). WWW '07. ACM, New York, NY, 271-280. DOI= <http://doi.acm.org/10.1145/1242572.1242610>
- [10] Frasincar, F., Borsje, J., Levering, L. 2009. A Semantic Web-Based Approach for Building Personalized News Services. *International Journal of E-Business Research* 5,3, 35-53.
- [11] Fu, X. 2007. Evaluating sources of implicit feedback in web searches. In *Proceedings of the 2007 ACM conference on Recommender systems* (Minneapolis, MN, USA, 2007). RecSys '07. ACM, New York, NY, 191-194. DOI= <http://doi.acm.org/10.1145/1297231.1297272>
- [12] Jonnalagedda, N., Gauch, S. 2013. Personalized News Recommendation Using Twitter. In *Proceedings of the 2013 IEEE/WIC/ACM International Joint Conferences on Web Intelligence (WI) and Intelligent Agent Technologies (IAT)* (Atlanta, GA, November 17-20, 2013), IEEE Press, New York, 3, 21-25. DOI= <http://dx.doi.org/10.1109/WI-IAT.2013.144>
- [13] Chen, T., Han, W.-L., Wang, H.-D., Zhou, Y.-X. 2007. Content Recommendation System Based on Private Dynamic User Profile. In *Proceedings of International Conference on Machine Learning and Cybernetics* (Hong Kong, August 19 - 22, 2007). IEEE Press, New York, 2112-2118.
- [14] Gabrilovich, E., Dumais, S., Horvitz, E. 2004. Newsjunkie: providing personalized newsfeeds via analysis of information novelty. In *Proceedings of the 13th international conference on World Wide Web* (New York, NY, USA, 2004). WWW '04. ACM, New York 482-490. DOI= <http://doi.acm.org/10.1145/988672.988738>
- [15] Goosen, F., Ijntema, W., Frasincar, F., Hogenboom, F. and Kaymak, U. (2011). News Personalization using the CF-IDF Semantic Recommender. In *Proceedings of WIMS'11* (Sogndal, Norway, May 25-27, 2011). WIMS '11. ACM, New York 1-12. DOI= <http://doi.acm.org/10.1145/1988688.1988701>
- [16] Gupta, A., Singh, K. 2013. Location based personalized restaurant recommendation system for mobile environments. In *Proceedings. on Advances in Computing, Communications and Informatics* (Mysore, August 22 - 25). IEEE Press, New York, 507-511. DOI= [10.1109/ICACCI.2013.6637223](http://doi.acm.org/10.1109/ICACCI.2013.6637223)
- [17] Hoffmann, P., Kaczmarek, A., Spaleniak, P., Kostek, B. 2014. Music Recommendation System. *Journal of Telecommunications and Information Technology*, 2, 59-69.
- [18] Lehtiniemi, A., Holm, J. 2014. Easy access to recommendation playlists: selecting music by exploring preview clips in album cover space. In *Proceedings of the 10th International Conference on Mobile and Ubiquitous Multimedia* (Beijing, China, 2011). MUM '11. ACM New York, 94-99.

- [19] IJntema, W., Goossen, F., Frasinicar, F. and Hogenboom, F. (2010). Ontology-based news recommendation. In *Proceedings of the 2010 EDBT/ICDT Workshops* (Lausanne, Switzerland, 2010). EDBT '10. ACM, New York, NY, 16:1-16:6. DOI= <http://doi.acm.org/10.1145/1754239.1754257>
- [20] Joachims, T., Freitag, D., Mitchell, T. 2007. WebWatcher: A Tour Guide for the World Wide Web. In *Proceedings of the International Joint Conference on Artificial Intelligence* (San Francisco, 2007). Morgan Kaufmann, 770–777.
- [21] Kompan, M., Bieliková, M. 2010. Content-Based News Recommendation. In *Lecture Notes in Business Information Processing*, Buccafurri, F., Semeraro, Ed. Springer, Heidelberg, 61, 61-72. DOI= 10.1007/978-3-642-15208-5_6
- [22] Lee, H. J., Park, S. J. 2007. MONERS: A news recommender for the mobile. *Expert Systems with Applications* 32, 1 (January 2007), 143-150. DOI= 10.1016/j.eswa.2005.11.010
- [23] Konstan, J. A., Miller, B. N., Maltz, D., Herlocker, J. L., Gordon, L. R., Riedl, J. 1997. GroupLens: applying collaborative filtering to Usenet news. *Commun. ACM* 40, 3 (March 1997), 77-87. DOI= <http://doi.acm.org/10.1145/245108.245126>
- [24] Li, L., Zheng, L., Yang, F., Li, T. 2014. Modeling and broadening temporal user interest in personalized news recommendation. *Expert Syst. Appl.* 41, 7 (June 2014), 3168-3177. DOI= <http://dx.doi.org/10.1016/j.eswa.2013.11.020>
- [25] Lv, Y., Moon, T., Kolari, P., Zheng, Z., Wang, X., Chang, Yi. 2011. Learning to model relatedness for news recommendation. In *Proceedings of the 20th international conference on World wide web* (Hyderabad, India, 2011). WWW '11. ACM New York, NY, 57-66. DOI= <http://doi.acm.org/10.1145/1963405.1963417>
- [26] Martínez, L., Rodríguez, R.M., Espinilla, M. 2009. REJA: A Georeferenced Hybrid Recommender System for Restaurants. In *Proceedings. International Joint Web Intelligence and Intelligent Agent Technologies* (Milan, Italy, September 15 – 18, 2009). IET, 87 – 190.
- [27] Barranco, M. J., Noguera, J. M., Castro, J., Martínez, L. 2013. Context-aware intelligent recommendation system for tourism. In *Proceedings. IEEE International Pervasive Computing and Communications Workshops* (San Diego, CA, March 18 – 22, 2013). IEEE Press, New York, 328-331. DOI= 10.1109/PerComW.2013.6529508
- [28] Montes-García, A., Álvarez-Rodríguez, J.M., Labra-Gayo, J.E. 2013. Towards a journalist-based news recommendation system: The Wesomender approach. *Expert Syst. Appl.* 40, 17 (December 2013), 6735-6741. <http://dx.doi.org/10.1016/j.eswa.2013.06.032>
- [29] Osinski, S. 2003. *An Algorithm for Clustering of Web Search Results*, Master thesis, Department of Computing Science, Poznań University of Technology.
- [30] Osinski, S., Weiss, D. User and Developer Manual for version 3.11.0-SNAPSHOT from <http://download.carrot2.org/head/manual/index.html>
- [31] Puolamäki, K., Salojärvi, J., Savia, E., Simola, J., Kaski, S. 2005. Combining eye movements and collaborative filtering for proactive information retrieval. In *Proceedings of the 28th annual international ACM SIGIR conference on Research and development in information retrieval* (Salvador, Brazil, 2005).SIGIR '05. ACM, New York, NY, 146-153. DOI= <http://doi.acm.org/10.1145/1076034.1076062>
- [32] Resnick, P., Iacovou, N., Suchak, M., Bergstrom, P., Riedl, J. 1994. GroupLens: an open architecture for collaborative filtering of netnews. In *Proceedings of the 1994 ACM conference on Computer supported cooperative work* (Chapel Hill, North Carolina, USA, 1994).CSCW '94. ACM, New York, NY, 175-186. DOI= <http://doi.acm.org/10.1145/192844.192905>
- [33] Schouten, K., Ruijgrok, P., Borsje, J., Frasinicar, F., Levering, L. and Hogenboom, F. (2010) A Semantic Web-based Approach for Personalizing News. In *Proceedings of the 2010 ACM Symposium on Applied Computing*. SAC '10. ACM, New York, NY 854-861. DOI=<http://doi.acm.org/10.1145/1774088.1774264>
- [34] Shapira, B., Shoval, P., Tractinsky, N., Meyer. 2009. epaper: A personalized mobile newspaper. *Journal of the American Society for Information Science and Technology*, 60, 11 (July 2009), 2333–2346. DOI= 10.1002/asi.21172
- [35] Soares, M., Viana, P. 2014. TV Recommendation and Personalization Systems: Integrating Broadcast and Video On demand Services, *Advances in Electrical and Computer Engineering* 14, 1 (February 2014), 115-120. DOI= 10.4316/AECE.2014.01018
- [36] Soares, M., Viana, P. 2015. Tuning metadata for better movie content-based recommendation systems. *Multimedia Tools and Applications*, 74, 17 (September 2015), 7015-7036. DOI= 10.1007/s11042-014-1950-1
- [37] Solomou, G. D., Kalou, A. K., Koutsomitropoulos, D. A. and Papatheodorou, T. S. (2011) A Mashup Personalization Service Based on Semantic Web Rules and Linked Data. In *2011 Seventh International Conference on Signal-Image Technology and Internet-Based Systems* (Dijon, November 28 - December 1, 2011). IEEE Press, New York 89–96. DOI=10.1109/SITIS.2011.12
- [38] Spink, A., Cole, C. 2015. New directions in cognitive information retrieval, 9:175, Springer, Netherlands.
- [39] Tavakolifard, M., Gulla, J. A. 2013. Tailored news in the palm of your hand: a multi-perspective transparent approach to news recommendation. In *Proceedings of the 22nd International Conference on World Wide Web* (Rio de Janeiro, Brazil, 2013). WWW '13. International World Wide Web Conferences Steering Committee, Republic and Canton of Geneva, Switzerland, 305-308. DOI= <http://dl.acm.org/citation.cfm?id=2487788.2487930>
- [40] Yang, F., Wang, Z.-M. 2009. A mobile location-based information recommendation system based on GPS and WEB2.0 services. *W. Trans. on Comp.* 8, 4 (April 2009), 725-734. DOI= <http://dl.acm.org/citation.cfm?id=1558756.1558770>
- [41] Reischach, F., Michahelles, F., Schmidt, A. 2009. The design space of ubiquitous product recommendation systems. In *Proceedings of the 8th International Conference on Mobile and Ubiquitous Multimedia* (Cambridge, United Kingdom, 2009). MUM '09. ACM, New York, NY, 2:1-2:10. DOI= <http://doi.acm.org/10.1145/1658550.1658552>