Friend Recommendation in a Social Bookmarking System: Design and Architecture Guidelines

Matteo Manca, Ludovico Boratto and Salvatore Carta

Abstract Social media systems allow users to share resources with the people connected to them. In order to handle the exponential growth of the content in these systems and of the amount of users that populate them, recommender systems have been introduced. As social media systems with different purposes arose, also different types of social recommender systems were developed in order to filter the specific information that each domain handles. A form of social media, known as *social bookmarking system*, allows to share bookmarks in a social network. A user adds as a friend or follows another user and receives updates on the bookmarks added by that user. In this paper, we present an analysis of the state-of-the-art on user recommendation in social environments and of the structure of a social bookmarking system, in order to derive design guidelines and an architecture of a friend recommender system in the social bookmarking domain. This study can be useful for future research, by highlighting the aspects that characterize this domain and the features that this type of recommender system has to offer.

Keywords Social bookmarking · Friend recommendation · User behavior · Tagging system

This work is partially funded by Regione Sardegna under project SocialGlue, through PIA—Pacchetti Integrati di Agevolazione "Industria Artigianato e Servizi" (annualità 2010).

M. Manca · L. Boratto (✉) · S. Carta
Dipartimento di Matematica e Informatica, Università di Cagliari,
Via Ospedale 72, 09124 Cagliari, Italy
e-mail: ludovico.boratto@unica.it

M. Manca
e-mail: matteo.manca@unica.it

S. Carta
e-mail: salvatore@unica.it

K. Arai et al. (eds.), *Intelligent Systems in Science and Information 2014*,
Studies in Computational Intelligence 591, DOI 10.1007/978-3-319-14654-6_14

1 Introduction

Social media systems are web-based services that allow users to build a public or semi-public profile, create a list of other users with whom they share a connection, and view and traverse their list of connections and those made by others within the system [5]. The widely-known and studied information overload problem, in these system took the name of "social interaction overload" [13, 27], which means that each user has to interact with an excessive amount of users and items. This leads to a scarcity of attention, which does not allow a user to focus on users or items that might be interesting for her/him. In order to face the social information overload problem, recommender systems have been adopted to filter the large amount of information available in the social domain; the class of recommender systems that operate in the social domain is known as *social recommender systems* [25]. These systems face the social interaction overload problem, by suggesting users or items that a target user might be interested in. In particular, user recommendation in a social domain aims at suggesting *friends* (i.e., recommendations are built for pairs of users that are likely to be interested in each other's content) or *people to follow* (i.e., recommendations are built for a user, in order to suggest users that might be interesting for her/him) [13].

User recommender systems that operate in the social media domain can be classified into three categories, based on the source of data used to build the recommendations:

1. Systems based on the analysis of social graphs, which explore the set of people connected to the target user in order to produce recommendations. These systems recommend either the closest users in the graph, like friends of friends and followees of followees (the "People you may know" feature offered by Facebook [24] is the most widely known example of this approach), or recommend the users that have the highest probability to be crossed in a random walk of the social graph (the main reference for this type of systems is the "Who to follow" recommendation in Twitter [12]).
2. Systems that analyze the interactions of the users with the content of the system (tags, likes, shares, posts, etc.). In order to exploit the user interests, these systems usually build a user profile by giving a structured form to content, thanks to the use of metrics like TF-IDF (Term Frequency—Inverse Document Frequency). Recommendations are produced by identifying users with similar profiles. An example of this class of systems is presented in Chen et al. [10].
3. Hybrid systems, that consider both the social graph and the interactions of the users with the content (an example is represented by [15]).

A *social bookmarking system* is a form of social media, which allows users to use keywords (*tags*) to describe resources that are of interest for them, helping to organize and share these resources with other users in the network [11]. The most widely-known examples of social bookmarking systems are Delicious,[1] where the

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

bookmarked resources are web pages, CiteULike,[2] where users bookmark academic papers, and Flickr,[3] where each picture can be annotated with tags.

Even if the use of these systems is widespread (in 2014, one million photos per day have been shared on Flickr[4]), to the best of the authors' knowledge, no approach in the literature recommended friends in a social bookmarking system prior to our recent works [20, 21].

In this paper we present a study that proposes the design and the definition of an architecture of a friend recommender system in a social bookmarking system. By analyzing the state-of-the-art on user recommendation in the social domain and how social bookmarking systems work, we design a friend recommender system that operates in this context and present its architecture.

The scientific contributions coming from this paper are the following:

- we analyze the state-of-the-art on user recommendation in social bookmarking systems, in order to highlight the weaknesses of the existing systems and derive the characteristics and features that a friend recommender system that operates in this domain has to offer;
- given the structure of a social bookmarking system and the analysis of the state-of-the-art, we present a design of a friend recommender system;
- we propose a novel architecture of a system to build friend recommendations in a social bookmarking system.

This paper extends the work presented in Manca et al. [20] in the following ways:

- a deeper contextualization with the state-of-the-art is going to be presented;
- the motivation to our study is going to be improved, by presenting an analysis of how our design guidelines relate to a real-world scenario. This will help us validate our study and introduce the architecture;
- an extension to the proposed architecture is provided, by presenting it at different granularities and by providing more details on each component. Moreover, we are going to analyze possible approaches to implement it in a real-world system and present possible extensions to it.

This study can be useful for any future research in this area, by presenting design guidelines and an architecture, which can be adopted in the development of a friend recommender system in the social bookmarking domain.

The rest of the paper is structured as follows: Sect. 2 presents the state-of-the-art on user recommendation in social environments; Sect. 3 illustrates how a social bookmarking system is structured and how it works; Sect. 4 presents the aspects related to the design of a friend recommender system in a social bookmarking system and presents guidelines, useful in the development of a system; Sect. 5 proposes an architecture of the system; Sect. 6 presents conclusions and future work.

2 http://www.citeulike.org/.

3 http://www.flickr.com/.

4 http://techcrunch.com/2014/02/10/flickr-at-10-1m-photos-shared-per-day-170-increase-since-making-1tb-free/.

2 Related Work

In the last years, social bookmarking systems have been studied from different points of view. This section presents related work on user recommendation in this research area. This study of the state-of-the-art will be deepened in Sect. 4, in order to analyze the aspects that characterize a recommender system that operates in this domain and the weaknesses of the existing approaches.

2.1 Systems Based on the Analysis of Social Graphs

In [12] authors present Twitter's user recommendation service, which allows to daily create a huge amount of connections between users that share common interests, connections and other factors. In order to perform the recommendations, the authors build a Twitter graph in which vertices represent users and the directed edges represent the "follow" relationship. The graph is stored in a graph database called FlockDB, and then data are processed with Cassovary (an open source in-memory graph processing engine). The system builds the recommendations by means of a user recommendation algorithm for directed graphs based on SALSA. In the next section, we are going to analyze this system, in order to design our proposal.

In [17] the authors model the user recommendation problem as a link prediction problem. They develop several approaches, that analyze the proximity of nodes in the graph of a social network, in order to infer the probability of new connections among users. Experiments show that the network topology is a good tool to predict future interactions.

In [2], Arru et al. propose a user recommender system for Twitter, based on signal processing techniques. The considered approach defines a pattern-based similarity function among users and makes use of a time dimension in the representation of the users profile. Our system is different, because we aim at suggesting friends while on Twitter there is no notion of "friend" but it works with "people to follow".

2.2 Systems Based on the Interactions with the Content

Quercia et al. [23] describe a user recommender system based on collocation. The proposed framework, called FriendSensing, recommends friends by analyzing collocation data. In order to produce the recommendations, the system uses geographical proximity and link prediction theories. In our domain we do not have such a type of information, so we cannot compare with this algorithm.

In [8], researchers present a study that considers different features in a user profile, behavior and network in order to explore the effect of *homophily* on user recommendations. They use the Dice coefficient on two users sets of tags and they find that similar tags do not represent a useful source of information for link prediction, while mutual followers are more useful for this purpose. As previously highlighted, the presented friend recommender system focuses on producing friend recommendation based on users' content (tag, bookmarks, etc.).

2.3 Hybrid Systems

In [29] authors propose a framework of user recommendation, based on users' interests and tested on Yahoo! Delicious. The proposed framework operates in two main steps: first, it models the users' interests by means of tag graph based community detection and represents them with a discrete topic distribution; then, it uses the Kullback-Leibler divergence function to compute the similarity between users' topic distribution and the similarity values are used to produce interest based user recommendation. Differently from this framework, the aim of the approach proposed in this paper is to produce friend recommendations (i.e., bidirectional connections) and not unidirectional user recommendations.

Chen et al. [10] present a people recommender system in an enterprise social network called Beehive, designed to help users to find known, offline contacts and discover new friends on social networking sites. With the proposed study, the authors demonstrate that algorithms that use similarity of user-created content were stronger in discovering new friends, while algorithms based on social network information were able to produce better recommendations.

In [15], the authors propose a user recommender system (called *Twittomender*) that, for each user, builds a user profile based on user's recent Twitter activity and user's social graph. The proposed system operates in two different manners; in the former mode the user puts a query and the system retrieves a ranking list of users, while in the latter mode the query is automatically generated by the system and it is mined by the user profile of the target user (the target user is the user that receives the recommendations). Our proposal does not use the social graph and, furthermore, in building recommendations it considers the friendship relationship and not the "user to follow" relationship.

In [14] authors present a recommender system for the IBM Fringe social network, based on aggregated enterprise information (like org chart relationships, paper and patent co-authorship, project co-membership, etc.) retrieved using SONAR, which is a system that allows to collect and aggregate these kinds of information. The authors deployed the people recommender system as a feature of the social network site and the results showed a highly significant impact on the number of connections on the site, as well as on the number of users who invite others to connect.

3 Social Bookmarking Systems

This section presents how a social bookmarking system is structured and how it works. This definition is based on the ones previously given in the literature (in particular we refer to [11, 16, 26]).

A social bookmarking system is composed by:

- a set of *users*;
- a set of *resources*. These resources characterize the type of social bookmarking system and, as mentioned in the introduction, they might be of different types (e.g., web pages);
- a set of *tags*, which are the keywords used to describe the resources;
- a set of *bookmarks*, which are represented as triplets (*user*, *resource*, *tag*); these triplets are known either as *tag assignments*, or as *tag applications*;
- a set of *connections* among users, which are represented as couples (*user*, *user*). Depending on the type of connection among two users, a couple might be ordered (i.e., users are connected by a *follow* relation), or not (i.e., users are *friends* and mutually follow each other). These connections form a graph, known either as *social graph* or *interest graph*.

Once a user decides to bookmark a resource by adding tags to it, these bookmarks are shown to the users who are friends with or follow this user.

Social bookmarking systems also offer privacy options, which allow to keep a bookmark private, or to share it only with a limited amount of users.

Features that allow to explore the tags and to facilitate the management of the bookmarks, like their export from browsers [19] and the possibility to add a bookmark to the profile by email, are often offered.

4 Designing a Friend Recommender System

The first objective of our proposal is to design a friend recommender system in a social bookmarking system. This section presents an analysis of the aspects that characterize both the state-of-the-art and social bookmarking systems, according to what was presented in the previous sections.

4.1 Analysis

In our analysis, we considered the following aspects:

(a) In [12], authors highlight that Twitter is an "interest graph", rather than a "social graph". A problem highlighted by the authors is that the analysis of such a graph suffers from scalability issues and, in order to contain the complexity of the recommender system, no user profile information could be used to build the recommendations. The definition of interest graph can also be extended to social bookmarking systems, since a user can add as a friend or follow another user, in order to receive her/his newly added bookmarks.

(b) Social media systems grow rapidly. This means that both the amount of content added to a social media system and the user population increase at a fast rate. A recommender system that operates in this context needs to build accurate user profiles, which are up-to-date with the constantly evolving preferences of the users.

(c) Resources usually have an unstructured form so, when building a content-based recommender systems, they are given a structured form, by introducing a *Content Analyzer* in the system [18].

(d) In the architecture of a content-based system, a *Feedback* component, which allows to update the user profile according to the recommended items that the user liked or did not like, is usually implemented [18].

(e) As [29] highlights, the tagging activity of the users reflects their interests. Therefore, the tags used by a user can be considered as an important source of information to exploit her/his interests.

Taking into account all these aspects, we drew the following conclusions.

Regarding point (a), in order to avoid the limitations related to the graph analysis in this domain, we aim at designing a system that only analyzes the content of the users (i.e., the tagged resources). So, we are going to design a system that belongs to the second class presented in the Introduction, i.e., the one that analyzes the interactions of the users with the content of the system.

Regarding points (b) and (c), given the rapid growth of information in social media systems, in order to efficiently and quickly update user profiles we decided to exploit the set of resources used by each user and the tags used to classify those resources, without using a *Content Analyzer* component, but analyzing only the *behavior of the users in the system*.

Regarding point (d), since the system we are designing deals with friend recommendations and we do not consider the connection between the users, the feedback of a user has no impact in her/his profile. On the contrary, when items are recommended in a content-based system, the feedbacks contain information about the preferences of the users, which help updating the user profiles.

Regarding point (e), we embraced the theory that user interests are reflected by the tagging activity and extended it, by following the intuition that users with similar interests use similar tags and the same resources.

4.2 Design Guidelines

Starting from the previous analysis, here we recap the features that a friend recommender system in the social bookmarking domain has to offer:

1. the resources saved by a user and the tags used to classify them represent a valuable source of information about a user. By monitoring them, we can constantly be updated on the interests of the users. Therefore, a friend recommender system in a social bookmarking system has to consider the tagged resources bookmarked by the users. Using only graph analysis to build the recommendations presents limitations, and building recommendations by

analyzing both the content the users interacts with and the interest graph would increase the complexity of the system (this might lead to the learning of user profiles that are not up-to-date with the current interests of the users);
2. the algorithms and metrics used by a system should be quickly computed, in order to keep the user profiles up-to-date. Therefore, we believe that a friend recommender system should mine *user behavior* (i.e., the interaction of the users with the content), more than the content itself. In fact, the introduction of a *Content Analyzer*, in order to give a structured form to the resources, would significantly increase the complexity of the system. In other words, it is harder to make an analysis of the content of each resource tagged by a user, instead of considering only the fact that a user is interested in that resource. Since social bookmarking systems grow at a fast rate, content analysis would lead to have outdated profiles and this component is discarded by our design and architecture;
3. in order to reduce the complexity of the system, and given the type of recommendations produced, the typical *Feedback* component of a Content-Based recommender system is removed when designing such a type of system. This choice was made since the accepted or rejected friends do not update the user profiles, which are built considering the tag assignments of the users;
4. in order to capture the interaction of the users on multiple levels and improve the capability to accurately recommend friends, a system has to be able to exploit all the sources of information coming from the tag assignments. Therefore, a friend recommender system has to analyze both the tags used by a user and the resources she/he bookmarked.

4.3 Design Guidelines Evaluation in a Real-World Scenario

In the following, an analysis of the user behavior in a social bookmarking system from a friend recommendation point of view is presented. In particular, how the bookmarking activity of a user is related to that of the others has been studied by analyzing a Delicious dataset, distributed for the HetRec 2011 workshop [9]. The dataset contains:

- 1,867 users;
- 69,226 URLs;
- 53,388 tags;
- 7,668 bi-directional user relations;
- 437,593 tag assignments [i.e., tuples (user, tag, URL)];
- 104,799 bookmarks [i.e., distinct pairs (user, URL)].

By analyzing user profiles, it emerges that users had an average of 123.697 tags, and an average of 56.132 bookmarked resources.

In order to be able to infer the possible connections among users, which might lead to friend recommendations, the number of common tags and resources between the users of the dataset have been computed, obtaining the following results: the average number of common tags among two users is 7.807, while the average number of common resources among two users is 0.042. In particular, considering only the users who have at least a common tag, the average number of common tags for a couple of users increases to 10.417; while considering only the users who have at least a common bookmarked resource, the average number of common resources for each couple of users increases to 1.673.

From the conducted analysis is possible to infer some properties related to the user behavior in a social bookmarking system, recapped below:

- the behavior of two users in a social bookmarking system is related both to the use of the tags and to the use of the resources;
- the use of tags represents a stronger form of connection (as also proved in the literature), with respect to the amount of common resources between two users. This happens because the probability that two users use the same tags is higher than the one to bookmark the same resource, since a user classifies a resource with more tags;
- by comparing the number of common tags and resources with respect to the number of all tags and resources, it emerges that the number of common tags and common resources is much smaller than the number of tags and resources used by each user (more precisely, 10.4 out of 123.7 tags, and 1.7 out of 56.1 resources).

This means that the behavioral analysis of a user, which characterized the design of the system, can be exploited in order to recommend friends in this domain. Therefore, following these guidelines, in the next section we are going to present a novel architecture to build friend recommendations by exploiting the behavior of the users in a social bookmarking system.

5 Architecture

In order to build an architecture for a friend recommender system in the social bookmarking domain, we are going to follow the design guidelines presented in the previous section. Figure 1 illustrates the high level view of the architecture.

While designing the system, in the first point of the guidelines we highlighted that we would only analyze the content of the system (i.e., the tag assignments). Therefore, the architecture does not have components that analyze the connections among users (i.e., who they follow or they are friends with).

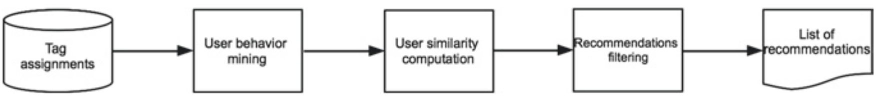

Fig. 1 High level architecture of the friend recommender system

The first task that the system has to compute is the mining of the user behavior by exploiting the tag assignments of each user (i.e., which resources a user tagged and with which tags). The *User behavior mining* component will allow to create a profile with the preferences of each user. Once the behavior of the user has been mined, a *User similarity computation* component will measure the similarity between the users. These similarities will then be inspected by the *Recommendation filtering* component, which will select the users most similar to each user, in order to recommend them.

The rest of the Section will provide the details of each high level component previously presented.

5.1 User Behavior Mining

As Fig. 2 shows, user behavior can be mined by two different components (i.e., the *Tag-based profile learner* and the *Resource-based profile learner*), following the considerations done on the fourth point of the design guidelines, which suggested to consider both the tags and the resources available in the bookmarks.

Taken as input the *Tag assignments* available for each user, two profile learner components will analyze the behavior of the user of her/his use of the tags (*Tag-based profile learner*) and on the bookmarked resources (*Resource-based profile learner*). Each component will now be presented in detail.

5.1.1 Tag-Based Profile Learner

Each time a user classifies a resource with a tag, her/his profile should be updated in order to capture the tagging behavior and build an accurate user profile. Taken as input the *Tag assignments* available for each user, this component builds a user profile, by considering the tags used by a user. Since in the design guidelines we highlighted the need to build profiles quickly, in order for them to be updated, this component might build profiles as binary vectors of the tags considered by users, or by considering the frequency of each tag used by a user. The output produced is a *Tag-based user profile*.

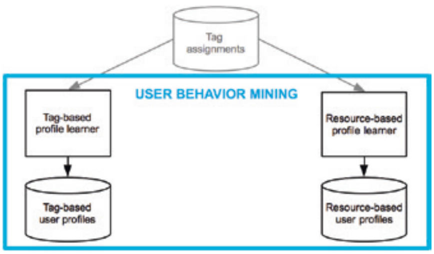

Fig. 2 Part of the architecture that mines user behavior to build the user profiles

5.1.2 Resource-Based Profile Learner

Given the *Tag assignments*, this component builds a second user profile, by analyzing the resources bookmarked by a user. Also this profile might be built as a vector, similarly to the possible implementations of the tag-based component. In case a binary vector is produced, it could highlight which resources have been bookmarked by the user and which not. Another possible implementation of this component would be by building a vector that contains in each element associated to a resource how many tags have been used to classify that resource (i.e., the relevance of a resource for a user could be measured by her/his effort to classify it, and a counter would keep track of this type of behavior). The output produced by this component is a *Resource-based user profile*.

5.2 User Similarity Computation

Figure 3 shows the two components that compute the similarities between the users, by comparing the tag-based user profiles (*Tag-based similarity computation* component) and the resource-based user profiles (*Resource-based similarity computation* component). This part of the architecture will now be described in detail.

5.2.1 Tag-Based Profile Similarity Computation

Given the *Tag-based user profiles* previously computed, this step estimates the association among each couple of tag-based user profiles, in order to derive how similar two users are. In case the similarity between binary vectors has to be computed, the Jaccard index would represent a standard measure to capture this similarity and efficient algorithms with low computational complexity have been proposed in the literature (e.g., the MinHash scheme [7], or the Signature scheme [1]). In case a vector with positive values is used to represent the profile, Pearson's correlation coefficient [22] as proved to be the most effective for the similarity assessment

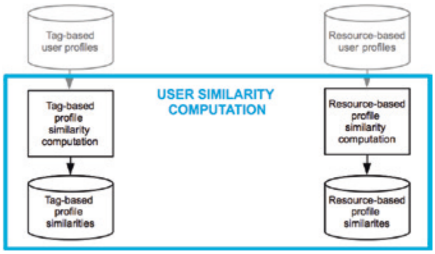

Fig. 3 Part of the architecture that computes the user similarities

between users [6]. Moreover, an efficient algorithm that exploits a support-based upper bound exists [28]. The output produced is a *Tag-based profile similarity*.

5.2.2 Resource-Based Profile Similarity Computation

Given the *Resource-based user profiles* previously computed, this step estimates the association among each couple of resource-based user profiles, in order to derive how similar two users are. According to the representation of the resource-based profile (i.e., binary vector or vector with positive values), the same algorithms used by the tag-based association component can be exploited. The output produced by this component is a *Resource-based profile similarity*.

5.3 Recommendations Filtering

This part of the architecture (shown in Fig. 4) solves the task of combining the tag-based and resource-based similarities between the users and filter them in order to produce the friend recommendations. Given the *Tag-based profile similarities* and *Resource-based profile similarities* previously built, the *Filtering component* selects the most similar users to recommend to the target user. For example, a threshold value might be used, in order to select only the users with high similarities with the target user. The output is a ranked *List of recommendations*, which contains the users to recommend to the target user.

5.4 System Architecture and Discussion

The full architecture of the system is presented in Fig. 5. Based on this structure and on the possible implementations previously presented, an efficient friend recommender system in the social bookmarking domain can be built.

Considering the social environment in which the recommendations have to be produced, an interesting aspect to notice in this architecture is that it lends itself

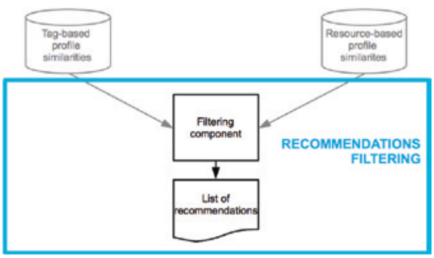

Fig. 4 Part of the architecture that produces the recommendations

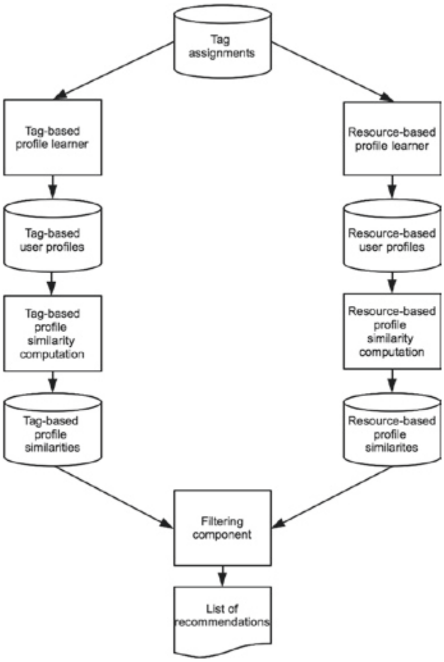

Fig. 5 Architecture of the friend recommender system

well to a parallelization on multiple machines. In fact, the two branches computed by the system (i.e., the one that operates with the tags and with the resources) can be independently computed.

Moreover, our architecture can be easily extended in case a new type of user behavior has to be mined. Suppose for example that we want to estimate the interest of the user on the topics of the resources.[5] A third branch could be added to this architecture, and this confirms that the proposed architecture can be adopted to build scalable systems. Given the possibility to extend our architecture to different types of behaviors to mine, this architecture can be used also to produce friend recommendations in different types of social media systems, by following the same pattern.

[5] Given that traditional techniques to manually categorize data cannot be applied in social environments [4] and that clustering techniques represent a good form to extract information for recommendation purposes [3], the resources could be clustered based on the tags used to classify them, in order to extract some meta-information about a group of resources related to a specific topic.

6 Conclusions and Future Work

This paper illustrated a study related to the design and the architecture of a friend recommender system in the social bookmarking domain. We analyzed the existing state-of-the-art works that recommend users in social domain and illustrated the structure of a social bookmarking system. This led to the design of a system that recommends friends in this context. After giving the design guidelines, the architecture of the system was presented. Following these design guidelines and this architecture, we built an efficient and very accurate friend recommender system [21], tested on the Delicious dataset previously illustrated. As future work we will implement the extension to the architecture proposed in the Discussion (i.e., the analysis of the topics) and the test it in our system.

References

1. Arasu, A., Ganti, V., Kaushik, R.: Efficient exact set-similarity joins. In: Proceedings of the 32nd International Conference on Very Large Data Bases. pp. 918–929. VLDB' 06, VLDB Endowment (2006). http://dl.acm.org/citation.cfm?id=1182635.1164206
2. Arru, G., Gurini, D.F., Gasparetti, F., Micarelli, A., Sansonetti, G.: Signal-based user recommendation on twitter. In: Carr, L., Laender, A.H.F., Lóscio, B.F., King, I., Fontoura, M., Vrandecic, D., Aroyo, L., de Oliveira, J.P.M., Lima, F., Wilde, E. (eds.) 22nd International World Wide Web Conference, WWW '13, Rio de Janeiro, Brazil, 13–17 May 2013, Companion Volume. pp. 941–944. International World Wide Web Conferences Steering Committee/ACM (2013)
3. Boratto, L., Carta, S., Manca, M., Mulas, F., Pilloni, P., Pinna, G., Vargiu, E.: A clustering approach for tag recommendation in social environments. Int. J. e-Bus. Dev. **3**, 126–136 (2013)
4. Boratto, L., Carta, S., Vargiu, E.: Ratc: A robust automated tag clustering technique. In: Noia, T.D., Buccafurri, F. (eds.) e-Commerce and Web Technologies. In: 10th International Conference, EC-Web 2009, Linz, Austria, 1–4 Sept 2009. Proceedings of Lecture Notes in Computer Science, vol. 5692, pp. 324–335. Springer (2009)
5. Boyd, D.M., Ellison, N.B.: Social network sites: Definition, history, and scholarship. J. Comput.-Mediated Commun. **13**(1), 210–230 (2007)
6. Breese, J.S., Heckerman, D., Kadie, C.: Empirical analysis of predictive algorithms for collaborative filtering. In: Proceedings of the Fourteenth conference on Uncertainty in artificial intelligence. pp. 43–52. UAI'98, Morgan Kaufmann Publishers Inc., San Francisco, CA, USA (1998). http://dl.acm.org/citation.cfm?id=2074094.2074100
7. Broder, A.: On the resemblance and containment of documents. In: Proceedings of the Compression and Complexity of Sequences 1997. pp. 21 SEQUENCES '97, IEEE Computer Society, Washington, DC, USA (1997). http://dl.acm.org/citation.cfm?id=829502.830043
8. Brzozowski, M.J., Romero, D.M.: Who should i follow? Recommending people in directed social networks. In: Adamic, L.A., Baeza-Yates, R.A., Counts, S. (eds.) Proceedings of the Fifth International Conference on Weblogs and Social Media, Barcelona, Catalonia, Spain, 17–21 July 2011. The AAAI Press (2011)
9. Cantador, I., Brusilovsky, P., Kuflik, T.: Second workshop on information heterogeneity and fusion in recommender systems (hetrec2011). In: Mobasher, B., Burke, R.D., Jannach, D., Adomavicius, G. (eds.) Proceedings of the 2011 ACM Conference on Recommender Systems, RecSys 2011, Chicago, IL, USA, 23–27 Oct 2011. pp. 387–388. ACM (2011)

10. Chen, J., Geyer, W., Dugan, C., Muller, M.J., Guy, I.: Make new friends, but keep the old: recommending people on social networking sites. In: Jr., DRO, Arthur, R.B., Hinckley, K., Morris, M.R., Hudson, S.E., Greenberg, S. (eds.) Proceedings of the 27th International Conference on Human Factors in Computing Systems, CHI 2009, Boston, MA, USA, 4–9 April 2009. pp. 201–210. ACM (2009)
11. Farooq, U., Kannampallil, T.G., Song, Y., Ganoe, C.H., Carroll, J.M., Giles, C.L.: Evaluating tagging behavior in social bookmarking systems: metrics and design heuristics. In: Gross, T., Inkpen, K. (eds.) Proceedings of the 2007 International ACM SIGGROUP Conference on Supporting Group Work, GROUP 2007, Sanibel Island, Florida, USA, 4–7 Nov 2007. pp. 351–360. ACM (2007)
12. Gupta, P., Goel, A., Lin, J., Sharma, A., Wang, D., Zadeh, R.: Wtf: the who to follow service at twitter. In: Schwabe, D., Almeida, V.A.F., Glaser, H., Baeza-Yates, R.A., Moon, S.B. (eds.) 22nd International World Wide Web Conference, WWW'13, Rio de Janeiro, Brazil, 13–17 May 2013. pp. 505–514. International World Wide Web Conferences Steering Committee/ACM (2013)
13. Guy, I., Chen, L., Zhou, M.X.: Introduction to the special section on social recommender systems. ACM TIST **4**(1), 7 (2013)
14. Guy, I., Ronen, I., Wilcox, E.: Do you know?: recommending people to invite into your social network. In: Conati, C., Bauer, M., Oliver, N., Weld, D.S. (eds.) In: Proceedings of the 2009 International Conference on Intelligent User Interfaces, 8–11 Feb 2009, Sanibel Island, Florida, USA. pp. 77–86. ACM (2009)
15. Hannon, J., Bennett, M., Smyth, B.: Recommending twitter users to follow using content and collaborative filtering approaches. In: Amatriain, X., Torrens, M., Resnick, P., Zanker, M. (eds.) Proceedings of the 2010 ACM Conference on Recommender Systems, RecSys 2010, Barcelona, Spain, 26–30 Sept 2010. pp. 199–206. ACM (2010)
16. Hotho, A., Jäschke, R., Schmitz, C., Stumme, G.: Bibsonomy: a social bookmark and publication sharing system. In: Proceedings of the First Conceptual Structures Tool Interoperability Workshop at the 14th International Conference on Conceptual Structures. pp. 87–102 (2006)
17. Liben-Nowell, D., Kleinberg, J.M.: The link prediction problem for social networks. In: Proceedings of the 2003 ACM CIKM International Conference on Information and Knowledge Management, New Orleans, Louisiana, USA, 2–8 Nov 2003. pp. 556–559. ACM (2003)
18. Lops, P., de Gemmis, M., Semeraro, G.: Content-based recommender systems: state of the art and trends. In: Ricci, F., Rokach, L., Shapira, B., Kantor, P.B. (eds.) Recommender Systems Handbook, pp. 73–105. Springer (2011)
19. Lund, B., Hammond, T., Hannay, T., Flack, M.: Social bookmarking tools (ii): a case study—connotea. D-Lib Magazine 11(4) (2005)
20. Manca, M., Boratto, L., Carta, S.: Design and architecture of a friend recommender system in the social bookmarking domain. In: Proceedings of the Science and Information Conference 2014. pp. 838–842 (2014)
21. Manca, M., Boratto, L., Carta, S.: Mining user behavior in a social bookmarking system—a delicious friend recommender system. In: Proceedings of the 3rd International Conference on Data Management Technologies and Applications (DATA 2014). pp. 331–338 (2014)
22. Pearson, K.: Mathematical contributions to the theory of evolution. iii. Regression, heredity and panmixia. Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Math. or Phys. Character (1896–1934) 187, 253–318 (1896)
23. Quercia, D., Capra, L.: Friendsensing: recommending friends using mobile phones. In: Bergman, L.D., Tuzhilin, A., Burke, R.D., Felfernig, A., Schmidt-Thieme, L. (eds.) Proceedings of the 2009 ACM Conference on Recommender Systems, RecSys 2009, New York, 23–25 Oct 2009. pp. 273–276. ACM (2009)
24. Ratiu, F.: Facebook: people you may know (May 2008), https://blog.facebook.com/blog.php?post=15610312130

25. Ricci, F., Rokach, L., Shapira, B.: Introduction to recommender systems handbook. In: Ricci, F., Rokach, L., Shapira, B., Kantor, P.B. (eds.) Recommender Systems Handbook, pp. 1–35. Springer, Berlin (2011)
26. Sen, S., Lam, S.K., Rashid, A.M., Cosley, D., Frankowski, D., Osterhouse, J., Harper, F.M., Riedl, J.: Tagging, communities, vocabulary, evolution. In: Hinds, P.J., Martin, D. (eds.) Proceedings of the 2006 ACM Conference on Computer Supported Cooperative Work, CSCW 2006, Banff, Alberta, Canada, 4–8 Nov 2006. pp. 181–190. ACM (2006)
27. Simon, H.A.: Designing organizations for an information rich world. In: Greenberger, M. (ed.) Computers, Communications, and the Public Interest, pp. 37–72. Johns Hopkins Press, Baltimore (1971)
28. Xiong, H., Shekhar, S., Tan, P.N., Kumar, V.: Exploiting a support-based upper bound of pearson's correlation coefficient for efficiently identifying strongly correlated pairs. In: Proceedings of the Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. pp. 334–343. KDD '04, ACM, New York, NY, USA (2004). http://doi.acm.org/10.1145/1014052.1014090
29. Zhou, T.C., Ma, H., Lyu, M.R., King, I.: Userrec: a user recommendation framework in social tagging systems. In: Fox, M., Poole, D. (eds.) Proceedings of the Twenty-Fourth AAAI Conference on Artificial Intelligence, AAAI 2010, Atlanta, Georgia, USA, 11–15 July 2010. AAAI Press (2010)