

FINDING DIACHRONIC LIKE-MINDED USERS

HOSSEIN FANI,^{1,2}  EBRAHIM BAGHERI,¹ FATTANE ZARRINKALAM,¹
XIN ZHAO,¹ AND WEICHANG DU²

¹Laboratory for Systems, Software and Semantics (LS3), Ryerson University, Toronto, ON, Canada

²Faculty of Computer Science, University of New Brunswick, Fredericton, NB, Canada

User communities in social networks are usually identified by considering explicit structural social connections between users. While such communities can reveal important information about their members such as family or friendship ties and geographical proximity, just to name a few, they do not necessarily succeed at pulling like-minded users that share the same interests together. Therefore, researchers have explored the topical similarity of social content to build like-minded communities of users. In this article, following the topic-based approaches, we are interested in identifying communities of users that share similar topical interests with similar *temporal* behavior. More specifically, we tackle the problem of identifying temporal (diachronic) topic-based communities, i.e., communities of users who have a similar temporal inclination toward emerging topics. To do so, we utilize multivariate time series analysis to model the contributions of each user toward emerging topics. Further, our modeling is completely agnostic to the underlying topic detection method. We extract topics of interest by employing seminal topic detection methods; one graph-based and two latent Dirichlet allocation-based methods. Through our experiments on Twitter data, we demonstrate the effectiveness of our proposed temporal topic-based community detection method in the context of news recommendation, user prediction, and document timestamp prediction applications, compared with the nontemporal as well as the state-of-the-art temporal approaches.

Received 13 September 2016; Revised 3 January 2017; Accepted 18 January 2017

Key words: community detection, time series analysis, topic detection.

1. INTRODUCTION

Social networks have shown to be an effective medium for communication and social interaction. Users can interact with others who share similar interests to communicate news, opinions, or other information of interest. As a result of such information sharing and communication processes, user communities emerge on social networks, which typically represent a group of like-minded or similarly behaving users (Natarajan et al. 2013; Abdelbary et al. 2014). Researchers have already investigated various techniques to identify and model communities within social networks to facilitate information flow and user connectivity (Sachan et al. 2012; Peng et al. 2015). Broadly speaking, there are two main approaches for community detection, namely, topology-based and topic-based techniques. Topology-based techniques consider explicit social connections between users (network structure), whereas topic-based approaches utilize information content posted by users (Ding 2011). Topology-based techniques may not be able to identify communities of users that share similar interests because of two reasons, among others, (i) there are many users on a social network that have similar interests but are not explicitly connected to each other, e.g., through follower or followee relationships, which is the primary requirement of topology-based techniques; (ii) many of the social connections may not be due to users' interest similarity but can be due to other factors such as friendship and kinship that do not necessarily point to inter-user interest similarity (Deng et al. 2013). Therefore, researchers have explored the possibility of utilizing the topical similarity of social content to build like-minded communities of users (Natarajan et al. 2013).

Address correspondence to Hossein Fani, Faculty of Computer Science, University of New Brunswick, Fredericton, NB, Canada; e-mail: hosseinfani@gmail.com

Most of the existing topic-based approaches define a community as a collection of users who share the same set of interests and do not necessarily interact with each other explicitly. Let us look at a concrete example from Twitter data in the last two months of the year 2010. By looking at the tweets, one can see that the two users @imadnaffa and @randytweety69 have been heavily engaged in posting content about the *WikiLeaks* event. These two users share the same topical interest, and topic-based approaches would consider them as the members of the same community. However, such approaches fall short when temporality of topical interests are of concern. In other words, topic-based approaches do not consider the temporal dimension of users' dispositions. In our example, @imadnaffa shows his interest toward the given topic in late November 2010, but @randytweety69 does not start posting about the same topic until much later in mid-December same year. As seen, while the users share the same topical interest, they do not show the same degree of interest toward it in similar time intervals. This distinction becomes important in applications such as news recommender systems. If the two mentioned users were placed in the same community, both users would obtain the same news recommendations on the *WikiLeaks* event, while @imadnaffa has already covered this event starting late November and may have moved on. As a result, she may not be interested in it any longer, while the topic would be of interest to @randytweety69. We believe and experimentally show in this article that for applications such as news recommendation, using communities that are sensitive to both topics and their temporal disposition are much more relevant.

The main objective of our work is to identify implicit user communities that have similar temporal dispositions with regard to similar topics of interest. Specifically, we want to identify those communities that distinguish between the users who are interested in a similar set of topics “this week,” e.g., @imadnaffa, from those who show the same behavioral pattern toward the same set of topics in the “following week,” e.g., @randytweety69; hence, supporting temporality in topic-based community detection methods. In this article, we propose a temporal topic-based community detection method based on multivariate time series analysis that measures inter-user similarity. Our approach is completely independent of the underlying topic detection method and is applicable in any textual content sharing network, which has timestamps for the shared content, e.g., tweets, blog posts, news articles, and citation networks, just to name a few.

Without loss of generality, we will focus on tweets in our work. We identify topics of interest within the Twitter social network using widely accepted topic detection methods, including two seminal probabilistic models, namely, latent Dirichlet allocation (LDA) (Blei et al. 2003) and topic over time (ToT) (Wang and McCallum 2006), and a graph-based approach (Weng and Lee 2011). Once the topics are identified, we model the users based on their temporal inclination toward topics using a multivariate time series representation. Our proposed user representation model captures both topic and time simultaneously and is completely independent of the underlying topic detection method. Based on the time series representation of the users, we are able to calculate user similarities and build a graph according to the measured user similarities. To find *temporal topic-based communities*, we apply graph clustering techniques on the user graph to extract cohesive subgraphs that would represent communities of users that are temporally and topically similar.

To illustrate the effectiveness of our proposed approach, we perform extensive experiments on a Twitter data set collected between November and December 2010. We adopt the evaluation strategy presented in Zarrinkalam et al. (2015) and Abel et al. (2011) and Wang and McCallum (2006) that consists of measuring the impact of our work on improving personalized news recommendation, user prediction, and document timestamp prediction.

The experimental results show that our proposed approach improves the performance of the recommendation and prediction applications.

The concrete contributions of our work are as follows:

- (1) We formally represent a user within temporal topic space through the use of multivariate time series. The proposed user representation effectively incorporates users' contributions toward the identified topics over time and is able to seamlessly integrate any topic detection methods and is, therefore, agnostic to the underlying topic detection method (Section 3.2.1).
- (2) We show how time series analysis techniques can be used to measure the similarity of pairs of users. This notion of similarity is further used to build a graph of user relations, not based on the users' social interactions, but rather based on their disposition toward similar topics in similar time intervals (Section 3.2.2).
- (3) We propose a graph representation of user interactions composed from their temporal and topical similarity and demonstrate how graph clustering can be used to identify user communities that consider both temporality and topical similarity when grouping users (Section 3.2.3).
- (4) We quantitatively demonstrate the effectiveness of our temporally detected communities in the context of applications such as news recommendation (Section 4.2.1), user prediction (Section 4.2.2), and document timestamp prediction (Section 4.2.3) in comparison with nontemporal as well as the state-of-the-art temporal community detection methods such as GrosToT (Hu et al. 2014).

The rest of the article is organized as follows: The next section reviews the related works after which the details of the proposed approach is introduced. Section 4 is dedicated to reporting our observations from our evaluations and experiments. Finally, Section 5 concludes the article and introduces future directions for research.

2. RELATED WORK

Existing user community detection approaches can be broadly classified into two categories (Ding 2011): topology-based and topic-based approaches. Topology-based community detection approaches represent the social network as a graph whose nodes are users and edges indicate explicit user relationships. This approach relies only on the graph structure of the social network and depends on notions such as components and cliques to extract communities (Fortunato 2010). On the other hand, topic-based approaches mainly focus on the information content of the users in the social network to detect latent communities. Because the goal of our proposed approach is to detect communities formed toward the topics extracted from users' information contents, we review topic-based community detection methods in this section.

Most of the works in topic-based community detection have focused on probabilistic models for detecting user communities based on textual content or jointly with social connections. For example, Abdelbary et al. (2014) have identified users' topics of interest and extracted latent communities based on the topics utilizing Gaussian restricted Boltzmann machines. Yin et al. (2012) have integrated community discovery with topic modeling in a unified generative model to detect communities of users who are coherent in both structural relationships and latent topics. In their framework, a community can be formed around multiple topics, and a topic can be shared among multiple communities. Sachan et al. (2012) have proposed probabilistic schemes that incorporate users' posts, social connections, and interaction types to discover latent user communities in social networks. In their work, they

have considered three types of interactions: a conventional tweet, a reply tweet, and a retweet. Other authors have also proposed variations of LDA, for example, author-topic model (Rosen-Zvi et al. 2004) and community-user-topic model (Zhou et al. 2006), to identify latent communities.

Another class of work attempts to transform the topic-based community detection problem into a graph clustering problem. These works are based on a similarity metric, which is able to compute the similarity of users based on their common topics of interest and a clustering algorithm to extract groups of users (latent communities) who have similar interests. For example, Liu et al. (2014) have proposed a clustering algorithm based on topic-distance between users to detect topic-based communities in a social tagging network. In this work, LDA is used to extract hidden topics in tags. Peng et al. (2015) have proposed a hierarchical clustering algorithm to detect latent communities from tweets. They have used the predefined categories in SINA Weibo and have calculated the pairwise similarity of users based on their degree of interest in each category.

None of the aforementioned methods incorporate temporal aspects of users' interests and undermine the fact that users of like-minded communities would ideally show similar contribution or interest patterns for similar topics throughout time. Hu et al. (2014) and Fani et al. (2016) are the only few that consider the notion of temporality. The authors in Hu et al. (2014) propose a unified probabilistic generative model, namely, GrosToT, to extract temporal topics and analyze topics' temporal dynamics in different communities. In our own previous work (Fani et al. 2016), we follow the same underlying hypothesis related to topics and temporality. However, our work distinguishes itself from Hu et al. (2014) in two ways. First, the authors in Hu et al. (2014) primarily aim at improving *topic detection* by introducing group-specific generative processes that are sensitive to time. In other words, their proposed generative model does not focus on finding better time-sensitive user groups but rather is aimed at identifying temporally sensitive topics. In contrast, the main purpose of our work is to exploit the temporal dynamics of users' behavior to enhance the identification of like-minded communities. Second, in contrast to GrosToT, we use time series analysis to model user's temporal dynamics. Our unique way of user representation provides us with the flexibility of being independent of any underlying topic detection methods, whereas GrosToT is primarily dependent on a variant of LDA for topic detection. This article extends our previous work (Fani et al. 2016) in the following respects: (i) We highlight that our approach is independent of the underlying topic detection method by adding a graph-based approach to our previous LDA-based approaches, and (ii) More comprehensive experiments are conducted and new findings are reported. Specifically, we introduce user prediction application to our evaluation strategies in addition to the previous applications of news recommendation and retweet prediction. Furthermore, we augment the experiments by adding GrosToT to the baselines as the state of the art in temporal community detection methods.

3. PROPOSED APPROACH

In our work, we aim at identifying latent temporal communities of users within a specific time period T , based on the temporal inclination of the users toward topics. We incorporate temporal aspects of users' interests and consider the fact that users of like-minded communities would ideally show similar contribution or interest patterns for similar topics throughout time. We divide this problem into two subproblems: *topic detection* and *community detection* in which the output of the first subproblem becomes the input of the second one. In this section, we concretely formulate these subproblems and propose our approach.

3.1. Topic Detection

Our proposed community detection method is able to seamlessly integrate any topic detection methods and is, therefore, agnostic to the underlying topic detection method. Hence, the focus of our work in this subproblem is not to propose a new topic detection method but rather to provide a common interface to the existing topic detection techniques for the purpose of temporal topic-based community detection. We highlight this by customizing one graph-based and two probabilistic LDA-based approaches in our work, as alternatives, to extract topics from documents. Foremost, we introduce the required preliminary definitions.

We model \mathbb{M} as the set of all documents where $\mathbb{M}_t^u \subset \mathbb{M}$ is the set of all documents posted at time interval t authored by user u . We represent the set of all distinct terms that have occurred in documents by \mathbb{W} , indexed from 1 to N . A document \mathbf{m} is a vector of N nonnegative integers, where the i^{th} element of the vector shows the occurrence frequency of the i^{th} term in that document.

We define topic \mathbf{z} to be a vector of N real numbers in $\mathfrak{R}^{[0,1]^N}$, summing to 1. The i^{th} number shows the participation score of the i^{th} term in forming the topic. Collectively, $\mathbb{Z} = \{\mathbf{z} \in \mathfrak{R}^{[0,1]^N} : \|\mathbf{z}\|_1 = 1\}$ is the set of all topics indexed from 1 to K . $\|\cdot\|_1$ is the L^1 -norm of \mathbf{z} . Topic distribution function $f : \mathbb{M} \rightarrow \mathfrak{R}^{[0,1]^K}$ maps a document to its topics, i.e., $\forall \mathbf{m} \in \mathbb{M} : f(\mathbf{m})[i]$ is the score of document \mathbf{m} with respect to topic \mathbf{z}_i such that $\|f(\mathbf{m})\|_1 = 1$. In the topic detection subproblem, given \mathbb{M} as input, we aim at identifying \mathbb{Z} , i.e. the topics formed in the documents posted in time period T .

Given \mathbb{M} , it is possible to extract topics \mathbb{Z} using various existing methods in the literature including topic detection methods introduced in Ding (2011), Blei et al. (2003), Zarrinkalam et al. (2015), and Weng and Lee (2011).

3.1.1. Graph-Based Approach. According to Weng and Lee (2011), one can utilize signal processing techniques to detect emerging topics. The fundamental hypothesis behind this topic detection method is that those terms who have correlated frequency within time could be considered to be conceptually related and can, therefore, collectively form a topic. To apply this approach, for any term $w \in \mathbb{W}$, a *term signal* is constructed. Simply, the term signal shows the number of times the term has been mentioned across all documents in different time intervals of time period T . More specifically, a term signal for term w is a temporally ordered set of integer values, expressed as $\mathbf{X}^w = (x_1^w, x_2^w, \dots, x_L^w)$, from discrete observations of term frequencies at L consecutive time intervals, such that x_t^w represents the occurrence number of the term w in all documents posted at time interval t .

We can calculate the similarity of two terms w_i and w_j , denoted by $d^{\mathbb{W}}(w_i, w_j)$, based on the cross-correlation of their term signals as follows:

$$d^{\mathbb{W}}(w_i, w_j) = \mathbf{X}^{w_i} \star \mathbf{X}^{w_j} = \sum_{-\infty}^{+\infty} (\mathbf{X}^{w_i})^* [m] \mathbf{X}^{w_j} [m] \quad (1)$$

where \mathbf{X}^w represent term signal for w , \star is the measure of cross-correlation between two term signals, and $(\mathbf{X}^w)^*$ is the complex conjugate of \mathbf{X}^w . Based on this, an undirected weighted term graph $\mathcal{G}^{\mathbb{W}} = (\mathbb{V}, \mathbb{E}, g)$ can be formulated such that $\mathbb{V} = \mathbb{W}$, $\mathbb{E} = \{e_{w_i, w_j} : \forall w_i, w_j \neq i \in \mathbb{W}\}$, and the weight function $g : \mathbb{E} \rightarrow \mathfrak{R}$ is defined as $g(e_{w_i, w_j}) = d^{\mathbb{W}}(w_i, w_j)$.

When the graph is constructed, graph partitioning algorithms such as the Louvain method (LM) (Blondel et al. 2008) can be used to identify highly cohesive subgraphs

(Weng and Lee 2011). Each subgraph represents an emerging topic on the text corpus at a given time period T . Here, each topic z is an induced subgraph \mathcal{G}^z of $\mathcal{G}^{\mathbb{W}}$ such that $\mathbb{V}^z \subseteq \mathbb{W}$, \mathcal{G}^z consists of all the edges of $\mathcal{G}^{\mathbb{W}}$ with incident vertices in \mathbb{V}^z , and $|\mathbb{V}^z| > 1$.

In accordance with our definition of topic $\mathbf{z} \in \mathbb{Z}$, we vectorize \mathcal{G}^z to N real numbers, summing to 1. To do so, for $1 \leq i \leq N$, we define the i^{th} number as the degree centrality of the term w_i if $w_i \in \mathbb{V}^z$ and 0 otherwise. Also, we normalize the result by its L^1 -norm. Finally, we define topic distribution function $f(\mathbf{m})[i] = \mathbf{m} \cdot \mathbf{z}_i$ where \mathbf{m} is a document, \cdot is the vector dot product, and $1 \leq i \leq K = |\mathbb{Z}|$.

3.1.2. Probabilistic Approaches. Latent Dirichlet allocation (LDA) assumes that a document is a mixture of topics and implicitly exploits co-occurrence patterns of terms to extract sets of correlated terms as topics of a text corpus (Blei et al. 2003). Similar to Hong and Davison (2010) and Weng et al. (2010), we see all terms extracted from documents of a user u for each time interval t , i.e., \mathbb{M}_t^u , as a single document $\mathbf{m} \in \mathbb{M}$. As another LDA-based approach, we use the topic over time (ToT) model (Wang and McCallum 2006), which simultaneously captures term co-occurrences and locality of those patterns over time and is hence able to discover more event-specific topics. In both LDA and ToT, $\mathbf{z} \in \mathbb{Z}$ is the multinomial distribution of terms specific to topic \mathbf{z} , and the topic distribution function f is defined as a Dirichlet distribution with parameter $\boldsymbol{\alpha}$; notationally, $f(\mathbf{m}) \sim \text{Dir}(\boldsymbol{\alpha})$.

After detecting topics \mathbb{Z} from a given document collection \mathbb{M} within a specific time period T and defining topic distribution function f using one of the aforementioned topic detection methods, our next goal is to identify communities of users formed on the basis of their temporal relation to the identified topics.

3.2. Temporal Topic-Based Community Detection

The core contribution of our work rests on identifying temporal topic-based communities. From a very abstract point of view, we model the subproblem of community detection on a set of users \mathbb{U} as a *set partitioning* problem. A partition \mathbb{P} of the set \mathbb{U} , $\mathbb{P}^{\mathbb{U}}$ is a set of nonempty subsets of \mathbb{U} as communities such that every user $u \in \mathbb{U}$ is in exactly one of these communities. The goal of community detection is to infer $\mathbb{P}^{\mathbb{U}}$ such that the users of high similarity be in the same community $\mathbb{C} \in \mathbb{P}^{\mathbb{U}}$, yet users of high dissimilarity be in different communities \mathbb{C}_i and $\mathbb{C}_{j \neq i}$. In our work, we consider two users to be similar if they share similar topics of interest and show similar temporal inclination toward them. Thus, our temporal topic-based community detection method seeks to find $\mathbb{P}^{\mathbb{U}}$ with respect to this sense of similarity, given the identified topics \mathbb{Z} from \mathbb{M} within a specific time period T and the topic distribution function f .

To do so, we represent the degree of contribution of a user to each topic $\mathbf{z} \in \mathbb{Z}$ over multiple time intervals as a vector. Collectively, this forms a multivariate time series for each user u toward all topics in \mathbb{Z} , which we refer to as the *user topic contribution time series*. We calculate the pairwise similarity between two users by computing the similarity between their corresponding user topic contribution time series. Based on these calculated similarities, we aim at calculating $\mathbb{P}^{\mathbb{U}}$. However, this would be considered to be a graph partitioning problem, which is NP-hard. Thus, we build a weighted graph of users and apply Louvain's heuristic in graph partitioning to detect user communities. Our approach for identifying temporal topic-based communities includes three steps: user representation, user similarity calculation, and user community identification, which are described in details as follows.

3.2.1. User Representation. We model each user's topics of interest and temporal inclination toward the topics through *user topic contribution time series*. Formally, the user

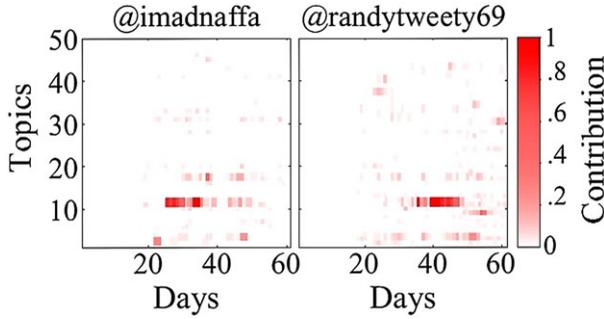


FIGURE 1. The heatmap of user topic contribution time series for two Twitter users.

topic contribution time series of user u for topic set \mathbb{Z} is a temporally ordered vectors of real values in L consecutive time intervals, expressed as $\mathbf{Y}^u = (\mathbf{y}_1^u, \mathbf{y}_2^u, \dots, \mathbf{y}_L^u)$. At each time interval t , \mathbf{y}_t^u is a vector whose i^{th} value $\mathbf{y}_t^u[i] \in \mathfrak{R}^{[0,1]}$ shows the degree of interest for the user u toward the topic \mathbf{z}_i . Assuming there are K topics detected, \mathbf{y}_t^u becomes a K -tuple vector, and the user topic contribution time series will be a K -variate time series. We apply the topic distribution function f of the topic detection method, as defined in 3.1, to each documents of user u at time interval t , \mathbb{M}_t^u , and sum over all with respect to each topics $\mathbf{z}_i \in \mathbb{Z}$; notationally, $\mathbf{y}_t^u[i] = \sum_{\mathbf{m} \in \mathbb{M}_t^u} f(\mathbf{m})[i]$.

In LDA and ToT topic models, we aggregate all documents of the user u in time interval t , \mathbb{M}_t^u , as a single document \mathbf{m} to find the topics. Thus, $\mathbf{y}_t^u = f(\mathbb{M}_t^u)$.

In Figure 1, we use heatmaps to visualize user topic contribution time series of two users from Twitter. In this figure, the Y -axis represents the topics indices, the X -axis denotes the time intervals enumerated for the last two months of the year 2010, and the density of the points show the contribution amplitude. The two users @imadnaffa and @randytweety69 mentioned in Figure 1 contributed to *WikiLeaks* (\mathbf{z}_{11} from Figure 8) but with time delays. User @imadnaffa posts about this topic in the end of November (day=30), whereas user @randytweety69 did not react to this topic till a week after (day=38).

The user topic contribution time series can be considered to be a good measure for finding the similarity between two users according to our definition of the latent user community. It allows finding like-minded users based on their *temporally* correlated contributions on similar topics. Based on Figure 1, nontemporal topic-based approaches group the two users, namely, @imadnaffa and @randytweety69, in the same community and consider them like-minded, because they are interested in the same topic, i.e., \mathbf{z}_{11} . However, the user @randytweety69 can be considered to be dissimilar from the other because the period of time during which she reacts to \mathbf{z}_{11} is not the same.

3.2.2. User Similarity. We compute the similarity of their corresponding user topic contribution time series to find the similarity of a pair of users. For this purpose, we employ the two-dimensional variation of the cross-correlation measure. The two-dimensional cross-correlation measure of two matrices $\mathbf{A}_{[C \times D]}$ and $\mathbf{B}_{[C \times D]}$, denoted by $\text{XC}_{[(2C-1) \times (2D-1)]}$, is calculated as follows:

$$\text{XC}[i, j](\mathbf{A}, \mathbf{B}) = \sum_{c=0}^{C-1} \sum_{d=0}^{D-1} \mathbf{A}[c, d] \mathbf{B}^*[c-i, d-j] \quad (2)$$

where \mathbf{B}^* denotes the complex conjugate of \mathbf{B} . Intuitively, the two-dimensional cross-correlation slides one matrix over the other and sums up the multiplications of the

overlapping elements. A positive row index i corresponds to a downward shift of the rows of \mathbf{A} over \mathbf{B} , and a negative column index j indicates a leftward shift of the columns. We can represent user topic contribution time series with respect to K number of topics in \mathbb{Z} in L consecutive time intervals as a $K \times L$ matrix. Then, the similarity of two users u_i and u_j , denoted as $usd^{\mathbb{U}}(u_i, u_j)$, can be defined based on the two-dimensional cross-correlation of their user topic contribution time series with no shift ($i = j = 0$), as follows:

$$usd^{\mathbb{U}}(u_i, u_j) = \frac{XC[0, 0](\mathbf{Y}^{u_i}, \mathbf{Y}^{u_j})}{\sqrt{(\mathbf{Y}^{u_i} \cdot \mathbf{Y}^{u_i})(\mathbf{Y}^{u_j} \cdot \mathbf{Y}^{u_j})}} \quad (3)$$

where \mathbf{Y}^u is the user topic contribution time series for user u .

We are now able to calculate the similarity between all pairs of users and group similar users that share similar temporal exposition toward similar topics of interest.

3.2.3. User Community. We identify user communities through graph-based partitioning heuristics. We represent users and their pairwise similarity through a weighted undirected graph. Precisely, let $\mathcal{UG}^{\mathbb{U}} = (\mathbb{V}, \mathbb{E}, g)$ be a weighted user graph in time period T such that $\mathbb{V} = \mathbb{U}$, $\mathbb{E} = \{e_{u_i, u_j} : \forall u_i, u_j \neq i \in \mathbb{U}\}$, and the weight function $g : \mathbb{E} \rightarrow \mathfrak{R}$ is defined as $g(e_{u_i, u_j}) = usd^{\mathbb{U}}(u_i, u_j)$. After constructing the user graph $\mathcal{UG}^{\mathbb{U}}$ for a given time period T , it is possible to employ a graph partitioning heuristic to extract partitions of users that form latent communities. As in graph-based topic detection, we leverage the Louvain method (LM). Louvain method is suitable for its following characteristics: (i) this algorithm can be applied to weighted graphs; (ii) it does not require a priori knowledge of the number of partitions when running the algorithm, and (iii) it is computationally very efficient when applied to large and dense graphs (Rotta and Noack 2011). While modularity maximization is NP-hard, the complexity of LM's greedy implementation is $O(n \log n)$, where n is the number of vertices (Blondel et al. 2008; Rotta and Noack 2011). Here, the output is a set of induced subgraphs of $\mathcal{UG}^{\mathbb{U}}$ representing temporal user communities $\mathbb{P}^{\mathbb{U}}$ that consist of like-minded users who have contributed to the same topics with the same temporal behavior and contribution degrees.

4. EXPERIMENTS

In this section, we describe our experiments in terms of the data set, setup, and comparative analytics. It should be noted that the main goal of our experiments is to determine the role and impact of temporality when building user communities. Therefore, we intentionally keep the parameters for topic detection methods constant (e.g., the number of topics in LDA and ToT) so as to avoid any unintended effects on the results and keep the scope of the experiments unchanged.

4.1. Data Set and Experimental Setup

In our experiments, we use a publicly available Twitter data set¹ collected and published by Abel et al. (2011). It consists of approximately 3M tweets posted by 135,731 unique users between November 1 and December 31, 2010. Each tweet, in addition to its text, includes a user ID and a timestamp.

¹<http://www.wis.ewi.tudelft.nl/umap2011>

Applying topic modeling methods such as LDA and ToT to extract topics from tweets might suffer from the sparsity problem (Sriram et al. 2010; Cheng et al. 2014), because they are designed for regular documents and not short, noisy, and informal texts like tweets. As suggested in Varga et al. (2014), to obtain better topics from Twitter without modifying the standard topic detection methods, we annotate each tweet $\mathbf{m} \in \mathbb{M}$ with concepts defined in Wikipedia using an existing semantic annotator. We see each concept as a term in the set \mathbb{W} . For instance, for a tweet such as “*Sweden issues Warrant for Wikileaks exec Julian Assange’s arrest http://bit.ly/9Ho0WM*,” a semantic annotator such as TagMe (Ferragina and Scaiella 2012) is able to identify and extract four Wikipedia concepts, namely, “Sweden,” “Arrest warrant,” “WikiLeaks,” and “Julian Assange.” Using concepts instead of words can lead to the reduction of noisy content within the topic detection process. Because each concept implicitly represents a collection of terms that are collectively more meaningful than a single word or a group of less coherent words (Ferragina and Scaiella 2012). As a result, the detected topics are more interpretable (Hulpus et al. 2013; Lau et al. 2011). We annotated the text of each tweet with Wikipedia concepts using the TagMe RESTful API², which resulted in 350,731 unique concepts. The choice of TagMe was motivated by a recent study that showed this semantic annotator performed reasonably well on different types of text such as tweets, queries, and web pages (Cornolti et al. 2013). We apply topic detection methods, described in Section 3.1, on the set of concepts extracted from the tweets to find topics \mathbb{Z} in our data set. The graph-based approach for topic detection (GbT) identifies topics by grouping a set of concepts that exhibit similar co-occurrence pattern over time. Given that our Twitter data set consists of tweets from a two-month period, we compute the pairwise similarities between daily ($L = 61$ days) concept signals. Because of the large number of identified concepts (350K), it is expensive to measure pairwise similarity through cross-correlation between all pairs of concepts. However, a large number of signals are trivial and not informative. We screen out the trivial concepts as suggested in Weng and Lee (2011). Filtering the trivial concepts significantly reduces the number of signals down to 782 and makes the computation of concept similarities practically feasible. The remaining concepts are then clustered using LM to form topics. We were able to find $K = 47$ topics, which served as our topic set \mathbb{Z}_{GbT} .

We also use LDA and ToT to discover topics. LDA-based approaches to topic detection need a priori knowledge of the number of topics, contrary to GbT. Therefore, we have opted to select the topic set size for LDA and ToT based on the number of topics detected by GbT. We aggregate daily tweets of each user to form a single document. Then, we apply LDA and ToT on the constructed documents to find topics, \mathbb{Z}_{LDA} and \mathbb{Z}_{ToT} , respectively. We have used MALLET³ for LDA and an open-source implementation available on GitHub⁴ for ToT.

Given the three extracted topic sets \mathbb{Z}_{GbT} , \mathbb{Z}_{LDA} , and \mathbb{Z}_{ToT} , we are interested in determining whether or not our temporal approach can provide a more accurate representation of user communities compared with the nontemporal and the state-of-the-art temporal approaches.

4.2. Evaluation Strategies and Results

We employ three *quantitative* evaluation strategies for evaluating our work. First, we adopt a widely used news recommendation approach to examine whether news recommendations based on temporal communities are more accurate (4.2.1. News Recommendation).

²<http://tagme.di.unipi.it/tag>

³<http://mallet.cs.umass.edu/topics.php>

⁴http://github.com/ahmaurya/topics_over_time

Second, we investigate whether the use of temporal communities can enhance the identification of the users who have posted news on the social network (4.2.2. User Prediction). Third, we compare the performance of our temporal approach to predict the timestamp of posted tweets (4.2.3. Timestamp Prediction). We perform the quantitative evaluations on our approach compare with two baselines: nontemporal and the state-of-the-art temporal approach GrosToT (Hu et al. 2014).⁵

To the best of our knowledge, GrosToT is the most related work to our objective on temporal community detection in the literature. GrosToT uses a unified probabilistic model to infer both topics and user communities together as latent factors. Assuming the number of topics K and the number of communities C are known in advance, it models temporal topic-based dynamism by multinomial distribution over time intervals for each topic and community. Specifically, it associates Dirichlet distributions for topics over words, communities over users, and topics over communities with different parameters, respectively. Also, a Dirichlet distribution for time is assigned given topic-community pairs. A user is a member of a community according to the assigned community-user distribution. Her tweet is generated based on multinomial distribution over, first, topic-community distribution to select topics and then topic-word distribution to select words. The timestamp of the tweet is obtained by the multinomial distribution of time-topic-community distribution. As seen, the model is based on the idea that there is a tight interrelation between communities and topics. This prevents its use where there is no community structure and impedes the integration of other topic detection methods for the task of community detection.

Additionally, we conclude this section by a *qualitative* account of the types and forms of communities built using our proposed approach. We practically show how communities of users have been separated based on their temporal inclination toward topics (4.2.4. Qualitative Analysis).

4.2.1. News Recommendation. To quantitatively evaluate the quality of our temporal topic-based communities, we deploy a typical news recommender application. Several researchers, such as Zarrinkalam et al. (2015) and Abel et al. (2011), have already suggested that the performance of user interest detection methods can be measured through observations made at the application level such as through news recommendation. Therefore, we opted to first evaluate our work based on such strategy but in a higher community level. To this end, we first build a gold standard data set by collecting news articles to which a user has explicitly linked in her tweets (or retweets). Our hypothesis is that users are interested in the topics of the news article, which they have posted about. Similar to tweets, we annotate news articles with Wikipedia concepts. A news article \mathbf{n} is a vector of N nonnegative integers, where the i^{th} number shows the occurrence frequency of the i^{th} concept and \mathbb{A} is the set of all news articles. We build the gold standard from a set of tweets that include a link to news article \mathbf{n} , posted by user u at time t . We drop the content of each tweet and save it as a triple (u, \mathbf{n}, t) consisting of the news article \mathbf{n} , user u , and the time t . As a result, $\mathbb{G} = \{(u, \mathbf{n}, t) : u \in \mathbb{U}, \mathbf{n} \in \mathbb{A}, 1 \leq t \leq L = 61\}$ forms our gold standard. It consists of 25,756 triples extracted from 3,468 distinct news articles posted by 1,922 users.

Given this gold standard, the objective is to see whether it would be possible to recommend the right news article to the users. A right news article \mathbf{n} to be recommended to a user u at time t would be the one that is included in the gold standard, that is, $(u, \mathbf{n}, t) \in \mathbb{G}$.

We build temporal topic-based communities according to our proposed approach for those users who have at least one triple, $\mathbb{U} = \{u : (u, \mathbf{n}, t) \in \mathbb{G}\}$. We create daily user

⁵ The implementation has been kindly provided by the authors.

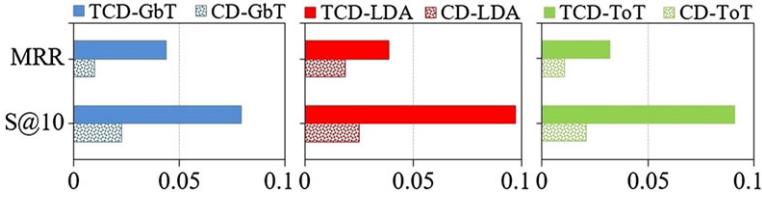


FIGURE 2. The performance of community detection methods in the context of news recommendation. TCD, temporal community detection method; CD, non-temporal community detection method; LDA, latent Dirichlet allocation; ToT, topic over time; MRR, mean reciprocal rank; GbT, graph-based approach for topic detection.

topic contribution time series for such users, i.e., $\mathbf{Y}^u = (\mathbf{y}_1^u, \mathbf{y}_2^u, \dots, \mathbf{y}_{L=61}^u)$, and compute the pairwise cross-correlation similarity on users' time series. Then, we build the weighted graph \mathcal{UG}^U and apply the LM by using its implementation from Pajek.⁶ This produces our temporal topic-based communities \mathbb{P}^U . We do these steps for each of our topic sets \mathbb{Z}_{GbT} , \mathbb{Z}_{LDA} , and \mathbb{Z}_{ToT} .

We also build nontemporal topic-based communities over the same set of users U . To do so, we project the daily user topic contribution time series of each user to the topic space by aggregating the values over the whole time period T . Formally, $\mathbf{y}_T^u = \sum_{t=1}^{L=61} \mathbf{y}_t^u$. Simply, \mathbf{y}_T^u is a vector which shows u 's degree of interest toward a set of topics in time period T . Then, we calculate the topic-based similarity of users u_i and u_j based on the cosine similarity of their corresponding $\mathbf{y}_T^{u_i}$ and $\mathbf{y}_T^{u_j}$. Finally, we create a weighted graph on the users and their similarity scores and apply the LM to find communities.

Because our main objective is not to propose a news recommender application, we adopt a simple recommender algorithm as follows: Given \mathbb{Z} , the set of K topics extracted in time period T , we represent each temporal user community $\mathbb{C} \in \mathbb{P}^U$ by a K -variate time series, named *community topic contribution time series*, and denoted as $\mathbf{Y}^{\mathbb{C}} = (\mathbf{y}_1^{\mathbb{C}}, \mathbf{y}_2^{\mathbb{C}}, \dots, \mathbf{y}_L^{\mathbb{C}})$ for L consecutive time intervals. $\mathbf{y}_t^{\mathbb{C}}$ represents \mathbb{C} 's contributions toward the topic set \mathbb{Z} at time t . Community topic contribution time series is calculated by aggregating the user topic contribution time series of all users who belong to user community \mathbb{C} at each L consecutive time intervals of T , i.e., $\mathbf{y}_t^{\mathbb{C}} = \sum_{u \in \mathbb{C}} \mathbf{y}_t^u$. We recommend news article \mathbf{n} at time t to a community \mathbb{C} according to the cosine similarity of the topic distribution of \mathbf{n} , i.e., $f(\mathbf{n})$, and $\mathbf{y}_t^{\mathbb{C}}$ which is the \mathbb{C} 's community topic contribution time series at time t .

In nontemporal communities, a user u has only \mathbf{y}_T^u , which shows u 's degree of interest toward the topics in the whole time period T . We build community-level degree of interest for the topics for each nontemporal community \mathbb{C} , denoted as $\mathbf{Y}_T^{\mathbb{C}}$, by summing over its members' topics of interest, i.e., $\mathbf{Y}_T^{\mathbb{C}} = \sum_{u \in \mathbb{C}} \mathbf{y}_T^u$. We recommend a news article \mathbf{n} to community \mathbb{C} based on the cosine similarity of topic distribution of \mathbf{n} , i.e., $f(\mathbf{n})$ and $\mathbf{Y}_T^{\mathbb{C}}$.

We use standard information retrieval metrics: mean reciprocal rank (MRR), which is the inverse of the first position that a correct item occurs within the ranked recommendations and success at rank k ($S@k$) that shows the probability that at least one correct item occurs within the top- k ranked recommendations. In the following, our approach and the two baselines are compared in terms of MRR and $S@10$. Compared with the nontemporal baseline, as shown in Figure 2, our temporal community detection method working on different topic detection methods, GbT, LDA, and ToT, outperforms the nontemporal counterparts in both metrics. This means that incorporating temporal aspects for extracting like-minded communities leads to more cohesive communities that consequently results in higher quality news

⁶mrvar.fdv.uni-lj.si/pajek

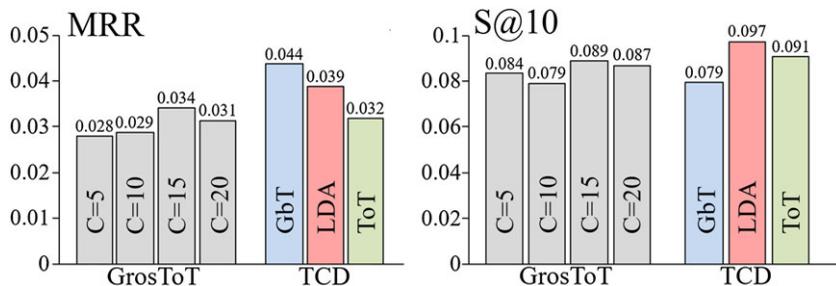


FIGURE 3. The performance of community detection methods in the context of news recommendation (TCD vs. GrosToT). TCD, temporal community detection method; GbT, graph-based approach for topic detection; LDA, latent Dirichlet allocation; ToT, topic over time; MRR, mean reciprocal rank.

recommendations. We will show in the qualitative evaluation (Section 4.2.4) that this translates into the desirable characteristic for our user communities that clearly separates users that have the same interests but in different time periods. This characteristic allows us to make recommendations at the appropriate time to users that are topically relevant.

To compare our approach with the state of the art, we run GrosToT on the ground truth. GrosToT requires prior knowledge about the number of topics and communities. We keep the number of topics constant and the same as for both LDA and ToT topic detection methods, i.e., $K = 50$, based on the reasoning provided in Section 4.1. We choose to increase the number of communities ($C \in \{5, 10, 15, 20\}$) for GrosToT till it shows no better performance. Also, GrosToT outputs a mixture according to a distribution that shows the degree of users’ membership to all communities. We partition the users by assigning a user to a community with the highest membership probability. Figure 3 depicts the performance of GrosToT as the number of communities changes compared with our proposed approach. Being independent of the underlying topic detection technique, our proposed temporal community detection method (TCD) is reported for three different topic detection techniques, namely, LDA, ToT, and GbT. As shown, all variants of TCD achieves consistently better or competitive performance compared with GrosToT where TCD-GbT and TCD-LDA show the best performance on MRR and S@10, respectively. The reason for this better performance by TCD could be the fact that the time series representation of the users captures both topical and temporal disposition of users more effectively and, consequently, the extracted user communities capture temporal topic-based similarity of users more coherently than GrosToT.

4.2.2. User Prediction. From the golden standard \mathbb{G} built in the previous section, we already know which users post a specific news article \mathbf{n} at time t . Based on this, given a news article \mathbf{n} , we are interested in predicting the users who have posted this news article at a specific time t . To identify such users, we determine those communities that show interest to topics of news article \mathbf{n} at time t . Our hypothesis is that the users who post this news article would be members of such communities. As we will show in the results, the predictions based on nontemporal communities do not seem to be accurate. An explanation can be that while some users may have contributed to the topics of the news article \mathbf{n} at time $t - 1$, but they may have shifted their interest as they progressed toward time t . This interest shifts would lead to poor performance in user prediction for nontemporal models.

To concretely perform the user prediction task, we iteratively consider the top- k news articles that belong to each of our communities. If the news article of interest is listed among the top- k relevant news for this community at iteration k , we consider the users in

TABLE 1. An instance of user prediction by news recommendation.

Observation	Top-k	\mathbb{C}_1 $\{u_1, u_2, u_3\}$	\mathbb{C}_2 $\{u_4, u_5, u_6\}$	Confusion matrix	Precision	Recall
$(\{u_1, u_2\}, \mathbf{n}_1, t_1)$	1	\mathbf{n}_1	\mathbf{n}_2	tp = 2, tn = 3 fp = 1, fn = 0	0.6	1.0
	2	\mathbf{n}_2	\mathbf{n}_1	tp = 2, tn = 0 fp = 4, fn = 0	0.3	1.0

Our observation shows that users u_1 and u_2 have posted the news article \mathbf{n}_1 in their tweets at time t_1 . We predict these users based on the members of communities \mathbb{C}_1 and \mathbb{C}_2 to which the news articles \mathbf{n}_1 and \mathbf{n}_2 are recommended for top-k; $k \in \{1, 2\}$.

that community to be potential posters of the news article. For the sake of clarifying this evaluation process, let us consider Table 1, which consists of two user communities \mathbb{C}_1 and \mathbb{C}_2 . Assume we are attempting to predict the user who posted news article \mathbf{n}_1 at time t_1 . At iteration 1, we identify the top 1 (the first) news article for each of the communities. If \mathbf{n}_1 is in top 1, then we have found a match, and the user who posted \mathbf{n}_1 is potentially a member of that community. In this case, the top 1 news for \mathbb{C}_1 is \mathbf{n}_1 ; therefore, we will predict that at iteration 1 (top 1), this news article is relevant for \mathbb{C}_1 , and hence, it could have been potentially posted by one of the members of $\mathbb{C}_1 = \{u_1, u_2, u_3\}$. From the observation, \mathbf{n}_1 has been posted by u_1 and u_2 at t_1 in reality. We also determined that the top 1 news for \mathbb{C}_2 at t_1 is \mathbf{n}_2 . Therefore, we estimate on this basis that $u_4, u_5, u_6 \in \mathbb{C}_2$ are not the posters of this news article. The confusion matrix can be developed based on these predictions. We then move to the second iteration, where we identify the top 2 news articles for each community at t_1 . We calculate the confusion matrix for this iteration based on whether \mathbf{n}_1 is one of the top 2 news articles of the communities. In Table 1, we can see that when considering the top 2 news articles, both \mathbb{C}_1 and \mathbb{C}_2 have \mathbf{n}_1 in their top 2. Our prediction would be that members of both \mathbb{C}_1 and \mathbb{C}_2 , i.e., $\{u_1, u_2, u_3\} \cup \{u_4, u_5, u_6\}$, are all possible posters of \mathbf{n}_1 leading to higher false positive rates.

We evaluate the quality of user predictions based on our temporal approach compared with the nontemporal communities. The results are presented in Figure 4 for top-k; $1 \leq k \leq 100$ in terms of precision and recall. As seen, our TCD under different topic detection methods, GbT, LDA, and ToT, unanimously outperforms the nontemporal counterparts in terms of both precision and recall. This reinforces the fact that our proposed approach produce communities that are topically and temporally coherent. When the poster of a certain news article needs to be identified at time t , both the content and time of the news can be taken into consideration, which would result in more accurate predictions. However, while nontemporal communities do consider the topic of the news article, they fail to take time into account and fall short in identifying changes in user interests. While a user may have had an interest in a certain topic in previous time intervals, she may have lost interest with time and therefore naturally be much less likely to post about that topic as time passes. For instance, let us consider the two sample Twitter users from Section 3.2. Both users @imadnaffa and @randytweety69 are interested in the *Wikileaks* topic (\mathbf{z}_{11} in Figure 8) but with a one week time difference. As was observed, @imadnaffa shows his interest in the topic toward late November, whereas @randytweety69 did so in mid-December. Now, if we observe a news article on Twitter talking about the *Wikileaks* topic on December 17, it is very likely that @randytweety69 is the user who is posting this news as opposed to the other user. The same logic applies if we see the same news article but this

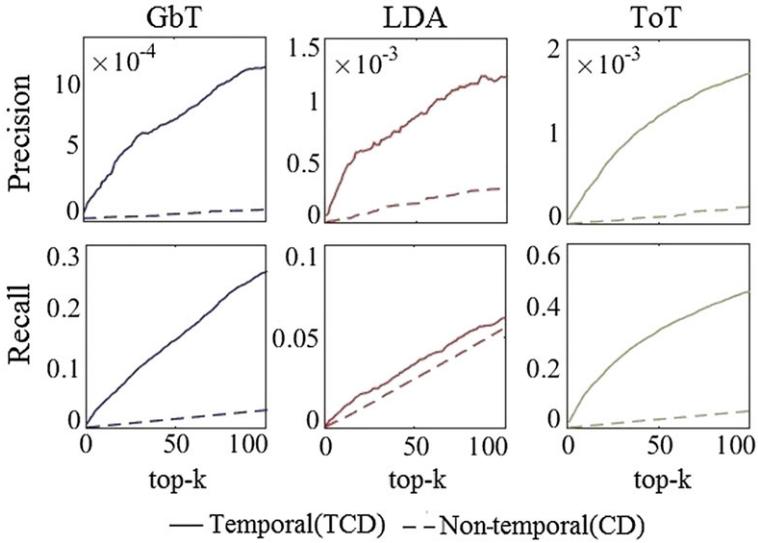


FIGURE 4. The results of the user prediction task (TCD: temporal vs. CD: nontemporal). TCD, temporal community detection method; GbT, graph-based approach for topic detection; LDA, latent Dirichlet allocation; ToT, topic over time.

time on November 25. This time, the likelihood of @imadnaffa posting this news is much higher. As it turns out in our experiments, the nontemporal community detection methods were not able to make a distinction between the two users and would hence predict both users to be the posters in both cases. This will result in many false positives, leading to a poor precision. However, temporal community detection in all three variations of topic sets \mathbb{Z}_{GbT} , \mathbb{Z}_{LDA} and, \mathbb{Z}_{ToT} were able to identify the correct user given the time and news article.

Moreover, in the ranked list of news articles for nontemporal communities, different news articles compete with each other only based on their topics, regardless of the temporal extension of topics in the community. Imagine community \mathbb{C} whose members have showed interest in two topics \mathbf{z}_i and \mathbf{z}_j , moving from \mathbf{z}_i toward \mathbf{z}_j with bursty behavior as the time passes from time interval t_{i-1} to t_i . The surge of posts from the members of the community on the second topic \mathbf{z}_j makes it the dominant topic for this community and undermines the existence of the first topic \mathbf{z}_i . Subsequently, we have news articles about the second topic \mathbf{z}_j with a higher rank than the first topic \mathbf{z}_i in the recommended list. As a result, it would become difficult to predict users who have posted content within this community on issues related to the first topic \mathbf{z}_i , producing more false negatives and lower recall in nontemporal communities. However, the ranked list of news articles to be recommended to a temporal community is in accordance with the community's topics of interest at each specific time interval. Back to our sample community \mathbb{C} , given a news article posted at time interval t_{i-1} about topic \mathbf{z}_i , it is highly probable that the temporal model will rank this news article higher in the recommended list for the community at time t_{i-1} (and lower at time t_i). This way, we can predict users who post about the first topic \mathbf{z}_i , which leads to less false negative and higher recall.

We evaluate the performance of the state-of-the-art competitor, i.e., GrosToT, in the task of user prediction as well. The results are presented in Figure 5 in terms of precision, recall, and f-measure. As observed, GrosToT with different number of communities does not show a coherent performance. While GrosToT with five communities ($C=5$) shows better performance in terms of recall compared with the other GrosToT variations, GrosToT with

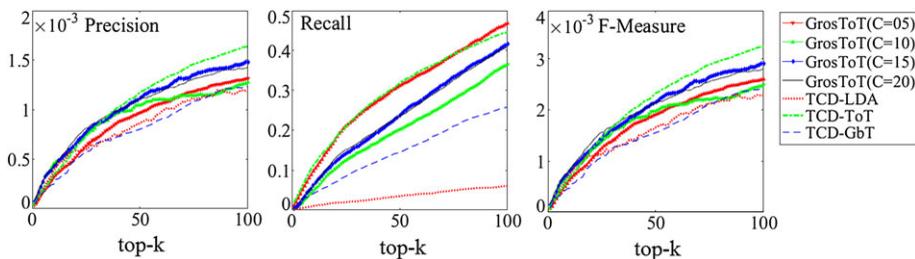


FIGURE 5. The results of the user prediction task (TCD vs. GrosToT). TCD, temporal community detection method; GbT, graph-based approach for topic detection; LDA, latent Dirichlet allocation; ToT, topic over time.

$C \in \{15, 20\}$ communities show higher precision. When comparing GrosToT with the best performing variation of our approach, i.e., TCD-ToT, one can make two observations: (1) TCD-ToT and GrosToT $_{C=5}$ show competitive performance in terms of recall. However, it should be noted that higher recall values for GrosToT $_{C=5}$ are expected given the fact that a lower number of communities will essentially group users in fewer clusters hence producing higher recall. However, when looking at the precision for GrosToT $_{C=5}$, it can be seen that this higher recall has come at the cost of a much lower precision compared with TCD-ToT. (2) In terms of precision, both TCD-ToT and GrosToT $_{C=20}$ show very competitive performance. GrosToT $_{C=20}$ performs slightly better for top- k ; $k \leq 35$, whereas TCD-ToT shows slightly better results for $k > 35$. Overall, when considering f-measure, GrosToT $_{C=20}$ and TCD-ToT show very competitive performance while TCD-ToT strongly outperforms its counterpart for top- k ; $k > 35$.

4.2.3. Timestamp Prediction. In this section, adopted from Wang and McCallum (2006), we evaluate our temporal communities in terms of the capability to predict the timestamp of documents. To do so, given a set of extracted user communities and a triple $(u, \mathbf{n}, t) \in \mathbb{G}$, where u is a user, \mathbf{n} is a news article, and t is the time the user posted the news article in her tweet, we want to predict t . For example, if two users @imadnaffa and @randytweety69 were posting about the topic *Wikileaks* (\mathbf{z}_{11}), we would be interested to know in which time intervals they posted about it. We refer to the predicted time as \hat{t} . To this end, knowing u , we first find the community to which u belongs. Then, within the selected community, the predicted \hat{t} is when the maximum cosine similarity of the community and the news article \mathbf{n} happens. For temporal communities, this can be performed by looking at the community topic contribution time series. Formally,

$$\text{maximize}_{1 \leq t \leq L} \left\{ \frac{\mathbf{y}_t^C \cdot f(\mathbf{n})}{|\mathbf{y}_t^C| |f(\mathbf{n})|} \right\}$$

where \cdot is the vector dot product, \mathbf{y}_t^C is the community topic contribution vector at time t , and $f(\mathbf{n})$ is the topic distribution function of the corresponding topic detection method applied on the news article \mathbf{n} .

Nontemporal communities do not have the time extension. To figure out when a nontemporal community reaches its peak for a topic, as mentioned in Wang and McCallum (2006), we build user topic contribution time series for its members and then community topic contribution time series the same way we do for temporal communities. We stress that we do build time series for nontemporal communities *after* the community detection. Users who share similar topics but with different temporal behavior would be members of the same nontemporal community. For instance, a nontemporal community of two users

@imadnaffa and @randytweety69 who are both interested in a topic on *Wikileaks* (\mathbf{z}_{11}) but in different time intervals, late November and mid-December, would have two peaks in the community topic contribution time series with respect to the topic \mathbf{z}_{11} . Therefore, there would be two possible predictions for \hat{t} in the triple $(@imadnaffa, \mathbf{n}, t) \in \mathbb{G}$, where \mathbf{n} is a news article about topic \mathbf{z}_{11} . Such situations lead to a poor time prediction as we will see in our experimentation.

Figure 6 compares temporal and nontemporal community detection methods based on their performance in timestamp prediction. The Y -axis shows the proportion of correct time predictions when the difference between the observed time t and the predicted time \hat{t} is less than the tolerance in X -axis.

As shown, temporal communities consistently outperform the nontemporal ones with respect to the tolerance range from the perfect match, i.e., same day, to the maximum possible, i.e., $L = 61$ days. From the figure, one can see that in TCD-GbT, we predict the time of mentions within 10 days with above 50% accuracy, while we have less than 44% in the CD-GbT for the same tolerance range. Similarly, for the TCD-ToT, we obtain more than 47% accuracy with 10 days margin of error, whereas CD-ToT gains 31% for the same error margin. A similar pattern can also be observed between TCD-LDA and CD-LDA where for the tolerance of 10 days, TCD-LDA offers 26% accuracy and CD-LDA only reaches 17%. It can be seen that regardless of the topic detection method, the temporal communities show a noticeable improvement over the nontemporal communities in terms of the accuracy of timestamp prediction.

Now, when comparing our method to the state-of-the-art temporal approach, GrosToT, it can be seen, as shown in Figure 7, that our methods TCD-GbT and TCD-ToT have a much higher accuracy in predicting the posting timestamps compared with all variations of the GrosToT method. This means that our proposed TCD partition users the best with regard to topics of interest and the users' respective temporal contributions, such that given an instance of user's topics of interest (e.g., news article posted by a user), the time of the user's contribution (timestamp of the news article) estimated by our proposed community of the user is the most accurate, comparing with the nontemporal methods (CD) and the state-of-the-art temporal approach (GrosToT).

Table 2 summarizes the results of our experiments. Overall, variants of our proposed TCD perform better than variants of GrosToT for both news recommendation and timestamp prediction applications. In the user prediction application, GrosToT and TCD show competitive performance. Beyond the comparison of the variants of TCD and GrosToT, one of the points that we would like to further elaborate on is the performance of the

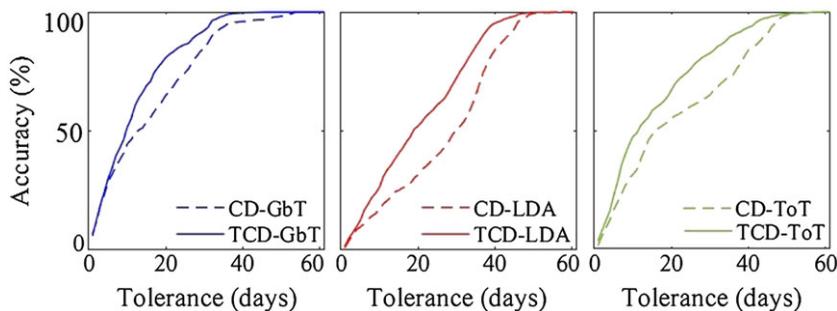


FIGURE 6. Time prediction accuracy (TCD: temporal vs. CD: nontemporal). TCD, temporal community detection method; GbT, graph-based approach for topic detection; LDA, latent Dirichlet allocation; ToT, topic over time.

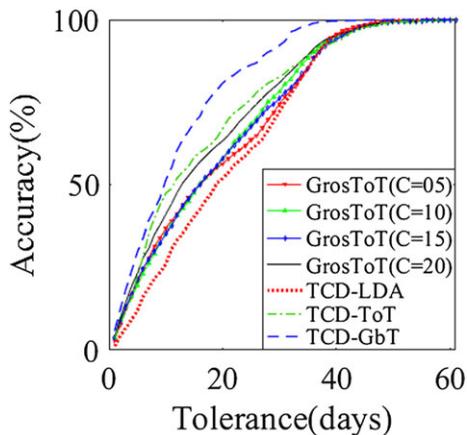


FIGURE 7. Time prediction accuracy (TCD vs. GrosToT). TCD, temporal community detection method; GbT, graph-based approach for topic detection; LDA, latent Dirichlet allocation; ToT, topic over time.

TABLE 2. Summary of experiments.

	News recommendation	User prediction	Timestamp prediction
Best performing model	TCD-LDA	$k \leq 35$: GrosToT $k > 35$: TCD-ToT	TCD-GbT

TCD, temporal community detection method; GbT, graph-based approach for topic detection; LDA, latent Dirichlet allocation; ToT, topic over time.

different variations of our own method. As discussed earlier, our proposed TCD approach is independent of the underlying topic detection method and can easily integrate any topic detection methods. To demonstrate this, we have reported results based on three different topic detection methods, namely, LDA, ToT, and GbT, in our experiments. In all three experiments, the performance of TCD when paired with different topic models differs depending on the application domain. This can be observed in Figures 3, 5, and 7, where TCD-LDA in news recommendation, TCD-ToT in user prediction, and TCD-GbT in the context of timestamp prediction have the best performance. This is in line with the discussion provided by Farzindar and Khreich (2015), which argues that event and topic detection methods need to be developed and adopted for a specific target application domain. In this respect, the flexibility to seamlessly integrate different topic modeling techniques seems an important advantage exhibited by our approach; TCD can be applied in a variety of application domains by pairing with the best topic detection method for that domain. However, this is not the case for the GrosToT that primarily relies on its integrated interdependence topic and user community modeling.

4.2.4. Qualitative Analysis. In this section, we intend to qualitatively corroborate the intuition that our communities are formed not only based on different topics of interest to the users but also based on the temporality of the user contributions. Without loss of generality and for discussion purposes, we demonstrate the behavior of our proposed work on TCD-GbT. We first show some of our identified topics in the GbT in Figure 8 along with their associated real-world events.

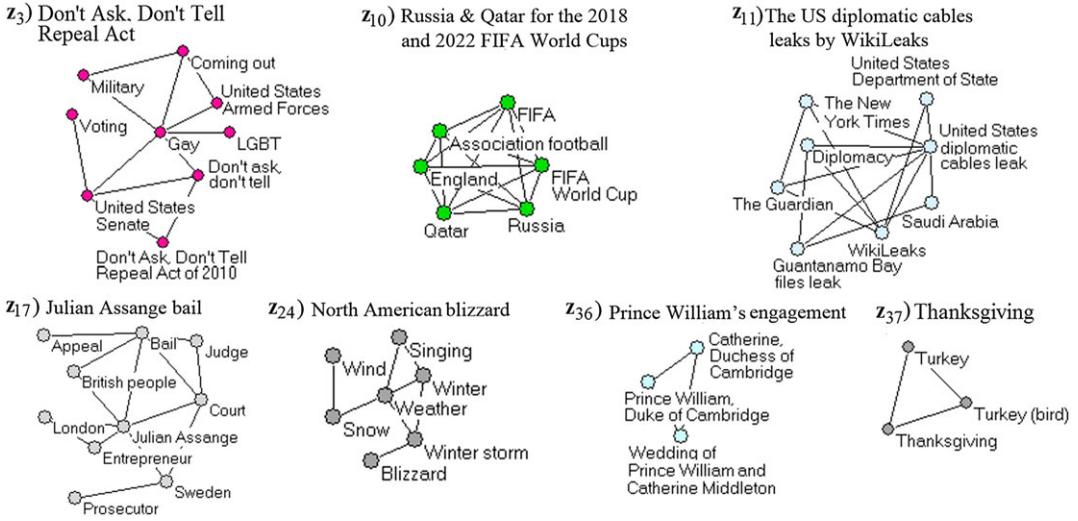


FIGURE 8. Sample topics from the graph-based topic detection method (GbT) that are relevant to the communities in Figure 9.

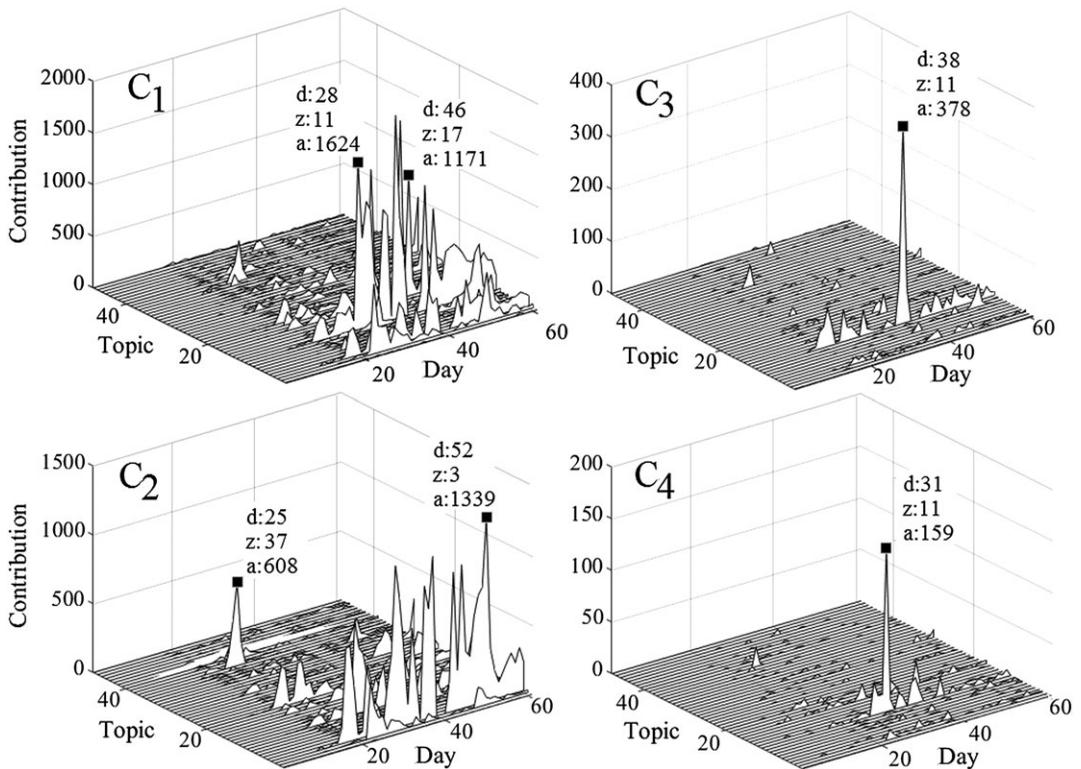


FIGURE 9. Sample temporal topic-based communities from our data set. The relevant topics to the communities in this figure are listed in Figure 8 (d , z , and a denote day, topic, and contribution amplitude).

We depict the temporal distribution of the topics over four of the extracted temporal topic-based communities from TCD-GbT in Figure 9. As shown, each community is illustrated in three dimensions of day (d), topic (z), and overall contribution amplitude (a), respectively. For instance, users in communities \mathbb{C}_1 and \mathbb{C}_2 discuss two disjoint sets of topics: *Julian Assange bail* (\mathbf{z}_{17}) and *WikiLeaks* (\mathbf{z}_{11}) in \mathbb{C}_1 and *Don't Ask, Don't Tell Repeal Act of 2010* (\mathbf{z}_3), and *Thanksgiving* (\mathbf{z}_{37}) in \mathbb{C}_2 . Here, the difference in topics forms different communities. However, the users of communities \mathbb{C}_3 and \mathbb{C}_4 discuss the same topic *WikiLeaks* (\mathbf{z}_{11}) but in different time intervals (with a week of delay). Nontemporal approaches would merge the users of such communities (\mathbb{C}_3 and \mathbb{C}_4) into a single community; however, our approach has been able to clearly distinguish between the users in these two communities. For instance, our approach ends with $@imadnaffa \in \mathbb{C}_4$ and $@randytweety69 \in \mathbb{C}_3$ for the two sample Twitter users in Figure 1. This is an important distinguishing feature for our work. As a case of news recommendation, it would be unreasonable to recommend a news article on topic \mathbf{z}_{11} to users in \mathbb{C}_4 on December 8, 2010 (day = 38) because the user had already discussed this topic 1 week ago on November 30, 2010 (day = 30). In contrary, it would make sense to recommend the same article to users of \mathbb{C}_3 who are actively pursuing the topic on Twitter on December 8, 2010.

5. CONCLUSION AND FUTURE WORK

In this article, we have proposed an approach to detect communities of like-minded users who share topics of interest with similar temporal behavior. We model the contribution of each user toward topics using multidimensional time series (3.2.1. User Representation) and apply two-dimensional cross-correlation on all pairs of such time series to find similar users in topics of interest and temporal behavior (3.2.2. User Similarity). We employ Louvain clustering, a heuristic graph partitioning algorithm based on modularity optimization, to create our final user communities (3.2.3. User Community). To find topics from the social network, we used state-of-the-art topic detection methods with different approaches, as alternatives, to show that our approach and its contribution is independent of topic detection algorithms. We used one graph-based and two probabilistic LDA and ToT methods in this article. We both *quantitatively* and *qualitatively* examine our approach. Our quantitative evaluation has been performed on three applications: news recommendation, user prediction, and timestamp prediction. According to our results, our temporal topic-based community detection method is able to effectively identify user communities that are formed around temporally similar behavior toward shared topics.

Possible future directions of our work would be as follows:

- (1) In our approach, we add the temporal dimension to topic-based communities. There are, however, methods in the literature that incorporate both topological and topical information. Integrating topological information in our approach when the underlying social network contains explicit ties among users would be an interesting direction for our future research.
- (2) In our experiments, it would be interesting to (i) consider different time intervals, e.g., weekly versus monthly, to construct the time series for the user representation; and (ii) employ overlapping clustering techniques for finding latent communities as opposed to only using partitioning techniques.
- (3) At a higher application level, the reasons behind the users' different temporal behavior toward shared topics still remain to be explored for comprehensive temporal modeling of the user community. Information diffusion and user interest preference may be two

promising areas for the observation of such effect that we would like to examine in our future work.

We believe that the work described in this article can be the foundation for future investigations of these and other important issues surrounding temporality in user community detection in social networks.

REFERENCES

- ABDELBARY, H. A., A. M. ELKORANY, and R. BAHGAT. 2014. Utilizing deep learning for content-based community detection. *In Science and Information Conference, London, UK*, pp. 777–784.
- ABEL, F., Q. GAO, G.-J. HOUBEN, and K. TAO. 2011. Analyzing user modeling on twitter for personalized news recommendations. *In User Modeling, Adaption and Personalization - Proceedings 19th International Conference, UMAP 2011, Girona, Spain*, pp. 1–12.
- BLEI, D. M., A. Y. NG, and M. I. JORDAN. 2003. Latent Dirichlet allocation. *Journal of Machine Learning Research*, **3**: 993–1022.
- BLONDEL, V. D., J.-L. GUILLAUME, R. LAMBIOTTE, and E. LEFEBVRE. 2008. Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, **2008**(10): P10008. <http://stacks.iop.org/1742-5468/2008/i=10/a=P10008>.
- CHENG, X., X. YAN, Y. LAN, and J. GUO. 2014. BTM: topic modeling over short texts. *IEEE Transactions on Knowledge and Data Engineering*, **26**(12): 2928–2941.
- CORNOLTI, M., P. FERRAGINA, and M. CIARAMITA. 2013. A framework for benchmarking entity-annotation systems. *In 22nd International World Wide Web Conference, WWW '13, Rio de Janeiro, Brazil*, pp. 249–260.
- DENG, Q., Z. LI, X. ZHANG, and J. XIA. 2013. Interaction-Based Social Relationship Type Identification in Microblog. *In Behavior and Social Computing, International Workshop on Behavior and Social Informatics, BSI 2013, Gold Coast, QLD, Australia, April 14-17, 2013 and International Workshop on Behavior and Social Informatics and Computing, BSIC 2013, Beijing, China, August 3-9, 2013, Revised Selected Papers*. Springer International Publishing, pp. 151–164.
- DING, Y. 2011. Community detection: topological vs. topical. *Journal of Informetrics*, **5**(4): 498–514.
- FANI, H., F. ZARRINKALAM, E. BAGHERI, and W. DU. 2016. Time-sensitive topic-based communities on twitter. *In Advances in Artificial Intelligence - Proceedings of the 29th Canadian Conference on Artificial Intelligence, Canadian AI 2016, Victoria, BC, Canada*, pp. 192–204.
- FARZINDAR, A., and W. KHREICH. 2015. A survey of techniques for event detection in twitter. *Computational Intelligence*, **31**(1): 132–164.
- FERRAGINA, P., and U. SCAIELLA. 2012. Fast and accurate annotation of short texts with wikipedia pages. *IEEE Software*, **29**(1): 70–75.
- FORTUNATO, S. 2010. Community detection in graphs. *Physics Reports*, **486**(3-5): 75–174.
- HONG, L., and B. D. DAVISON. 2010. Empirical study of topic modeling in twitter. *In Proceedings of the First Workshop on Social Media Analytics, SOMA '10. ACM, New York, NY, USA*, pp. 80–88.
- HU, Z., J. YAO, and B. CUI. 2014. User group oriented temporal dynamics exploration. *In Proceedings of the Twenty-Eighth AAAI Conference on Artificial Intelligence, Québec City, Québec, Canada*, pp. 66–72.
- HULPUS, I., C HAYES, M KARNSTEDT, and D GREENE. 2013. Unsupervised graph-based topic labelling using dbpedia. *In Sixth ACM International Conference on Web Search and Data Mining, WSDM 2013, Rome, Italy*, pp. 465–474.
- LAU, J. H., K. GRIESER, D. NEWMAN, and T. BALDWIN. 2011. Automatic labelling of topic models. *In The 49th Annual Meeting of the Association for Computational linguistics: Human Language Technologies, Portland, Oregon, USA*, pp. 1536–1545.

- LIU, H., H. CHEN, M. LIN, and Y. WU. 2014. Community detection based on topic distance in social tagging networks. *TELKOMNIKA Indonesian Journal of Electrical Engineering*, **12**(5): 4038–4049.
- NATARAJAN, N., P. SEN, and V. CHAOJI. 2013. Community detection in content-sharing social networks. *In Proceedings of the 2013 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, ASONAM '13*. ACM: New York, pp. 82–89.
- PENG, D., X. LEI, and T. HUANG. 2015. DICH: a framework for discovering implicit communities hidden in tweets. *World Wide Web*, **18**(4): 795–818.
- ROSEN-ZVI, M., T. L. GRIFFITHS, M. STEYVERS, and P. SMYTH. 2004. The author-topic model for authors and documents. *In UAI '04, Proceedings of the 20th Conference in Uncertainty in Artificial Intelligence, Banff, Canada*, pp. 487–494.
- ROTTA, R., and A. NOACK. 2011. Multilevel local search algorithms for modularity clustering. *ACM Journal of Experimental Algorithmics*, **16**: 2.3:2.1–2.3:2.27.
- SACHAN, M., D. CONTRACTOR, T. A. FARUQUE, and L. V. SUBRAMANIAM. 2012. Using content and interactions for discovering communities in social networks. *In Proceedings of the 21st World Wide Web Conference, WWW 2012, Lyon, France*, pp. 331–340.
- SRIRAM, B., D. FUHRY, E. DEMIR, H. FERHATOSMANOGLU, and M. DEMIRBAS. 2010. Short text classification in twitter to improve information filtering. *In Proceedings of the 33rd International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR '10*, pp. 841–842.
- VARGA, A., A. E. C. BASAVE, M. ROWE, F. CIRAVEGNA, and Y. HE. 2014. Linked knowledge sources for topic classification of microposts: a semantic graph-based approach. *Journal of Web Semantics*, **26**: 36–57.
- WANG, X., and A. MCCALLUM. 2006. Topics over time: a non-Markov continuous-time model of topical trends. *In ACM SIGKDD*.
- WENG, J., and B.-S. LEE. 2011. Event detection in twitter. *In Proceedings of the Fifth International Conference on Weblogs and Social Media, Barcelona, Catalonia, Spain*.
- WENG, J., E.-P. LIM, J. JIANG, and Q. HE. 2010. Twitterrank: finding topic-sensitive influential twitterers. *In Proceedings of the Third International Conference on Web Search and Web Data Mining, WSDM 2010, New York, NY, USA*, pp. 261–270.
- YIN, Z., L. CAO, Q. GU, and J. HAN. 2012. Latent community topic analysis: integration of community discovery with topic modeling. *ACM TIST*, **3**(4): 63:1–63:21.
- ZARRINKALAM, F., H. FANI, E. BAGHERI, M. KAHANI, and W. DU. 2015. Semantics-enabled user interest detection from twitter. *In IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology, WI-IAT 2015, Singapore, Volume I*, pp. 469–476.
- ZHOU, D., E. MANAVOGLU, J. LI, C. L. GILES, and H. ZHA. 2006. Probabilistic models for discovering e-communities. *In Proceedings of the 15th International Conference on World Wide Web, WWW 2006, Edinburgh, Scotland, UK*, pp. 173–182.