

Research Article

Extraction of Multilayered Social Networks from Activity Data

Katarzyna Musiał,¹ Piotr Bródka,² Przemysław Kazienko,² and Jarosław Gaworecki³

¹ School of Natural and Mathematical Sciences, Department of Informatics, King's College London, London WC2R 2LS, UK

² Institute of Informatics, Wrocław University of Technology, Wyb. Wyspiańskiego 27, 50-370 Wrocław, Poland

³ Research & Engineering Center Sp. z o.o., Ulica Strzegomska 46B, 53-611 Wrocław, Poland

Correspondence should be addressed to Przemysław Kazienko; kazienko@pwr.wroc.pl

Received 23 January 2014; Accepted 6 March 2014; Published 2 July 2014

Academic Editors: H. R. Karimi, Z. Yu, and W. Zhang

Copyright © 2014 Katarzyna Musiał et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The data gathered in all kinds of web-based systems, which enable users to interact with each other, provides an opportunity to extract social networks that consist of people and relationships between them. The emerging structures are very complex due to the number and type of discovered connections. In web-based systems, the characteristic element of each interaction between users is that there is always an object that serves as a communication medium. This can be, for example, an e-mail sent from one user to another or post at the forum authored by one user and commented on by others. Based on these objects and activities that users perform towards them, different kinds of relationships can be identified and extracted. Additional challenge arises from the fact that hierarchies can exist between objects; for example, a forum consists of one or more groups of topics, and each of them contains topics that finally include posts. In this paper, we propose a new method for creation of multilayered social network based on the data about users activities towards different types of objects between which the hierarchy exists. Due to the flattening, preprocessing procedure of new layers and new relationships in the multilayered social network can be identified and analysed.

1. Introduction

Nowadays, for the first time, we have possibility to process big data about interactions and activities of millions of individuals gathered in all sorts of web-based systems. Communication technologies allow us to form large networks, which in turn shape and catalyse our activities. Due to their scale, complexity, and dynamics, these networks are extremely difficult (or impossible) to analyse in terms of traditional social network analysis methods. The analysis of the network data is at the very early stages and requires a lot of efforts in both developing tools and approaches to tackle it as well as understanding the nature and functioning of networks extracted from this data. The process of network creation is not as straightforward as it seems to be. In the web-based systems, users can interact with each other via different communication channels and utilize various services. This implies that the relationships between users can be extracted based on both direct and indirect communications. The former one is, for example, sending emails or video calls

where information is passed directly from one person (group of people) to other(s), whereas to the latter we can count in, for example, commenting objects in the multimedia sharing systems or using the same tags to describe the objects. In both situations there are objects (e.g., email, photo, and tag) that serve as medium in communication between users. The types of these objects differ depending on the web system; for example, at the Internet forum the objects are groups of topics, topics, and posts while in the email service it will be a single message. Additionally, within a single system, the hierarchies of objects can exist.

Extraction of a social network in the environment where users interact with each other using different objects, which create hierarchies, is the main contribution of this paper. In order to perform this task first the hierarchical presocial network (HPSN), where relation between users and objects and between objects exists (Section 5), must be created. After that, the flattening process, in which the hierarchy of objects is removed, is performed (Section 6). As a result, the flat presocial network (FPSN) is obtained where the only

connections that exist are between users and one type of the previously chosen objects. Based on FPSN, social network, where only the connections between users exist, is created (Section 7). The whole idea is presented using simple case study of the Internet forum (Section 8). Finally, the real-world experiments were performed and their outcomes are presented in Section 9.

2. Related Work

There are many types of complex network systems. One of the classifications distinguishes between infrastructures and natural complex systems [1]. The former are physical systems (energy and transportation networks) and virtual systems (Internet, WWW, and telecommunication), whereas the latter are biological networks, social networks, food webs, and ecosystems.

The type of complex systems that is investigated in this paper is a social network formed by people who interact with each other or take part in common activities. The concept of social network has been described by different researchers [2–6] and the definition that is commonly and widely used is that a *social network* is a finite set of individuals, who are the nodes of the network and activities or relations between them, which are represented by edges of the network. A social network (SN) commonly represents the mutual communication and/or activity occurring between users as well as their direction, intensity, and profile [7].

It should be noted that, during analysis of social networks, researchers usually take into account only one activity type while in most cases many different types of relationships exist between users. The special type of social networks that allows the representation of many different activities is called a multilayered social network [5, 8–15], a.k.a. multidimensional network [16–18], or multiplex networks [7, 19–22]. Even the same authors use different names for this kind of complex networks; compare, for example, [12, 21]. Overall, due to high complexity, such networks are more difficult to be extracted and analysed than simple one-layered networks.

Sociologists and psychologist typically create questionnaires and perform interviews in order to collect data which allow them to create and analyse social networks. However, nowadays, the rapid development of the Internet and telecommunication together with the ease of gathering vast amount of data have created the possibility for IT systems to provide vast amount of information about users activities. As a result, researchers have now easy access to big datasets about people's activities ready to analyse. Social networks can be extracted from, for example, bibliographic data [23], blogs [24], photos sharing systems like Flickr [25], e-mail systems [26], telecommunication data [27, 28], social services like Twitter [29] or Facebook [30, 31], video sharing systems like YouTube [32], Wikipedia [33], and much more. Moreover, the whole separate systems were created only for the extraction, aggregation, and visualization of social networks [34, 35].

Nevertheless, as mentioned before, only few scientists have focused their research interests on multilayer social network extraction from activity data [14, 16, 18, 20, 22, 36–39].



FIGURE 1: The object-based relationship in the social network on the Internet.

Moreover, no one has studied the hierarchy and relationships between objects in this data. The only hierarchical dependencies in the social networks that were analysed were associated with the hierarchy between users such as employee-employer and the employee-manager [40]. Thus, the analysis of hierarchy between objects presented in this paper is a new approach to extract multilayer social networks from the activity data.

3. Object-Based Relationships

In web-based social systems, there is always an object that plays a role of “a middleman” in a relationship between two users (Figure 1) [41]. In the case of direct communication, people send e-mails to each other or make videoconferences or phone calls via VoIP services. In those cases all participants are aware of the existing relationship. However, sometimes two users can be in a relationship but they do not maintain it actively and consciously, for example, people who comment on the same blog or participate in the same conference. These types of common activities can result in indirect relationships. The roles of both users in an indirect relationship towards the object can be either the same or different (Figure 2).

Object-based relation with equal roles r_{xy}^{aa} means that users x and y meet each other through the object and their role in relation to this object is the same. In other words, they participate in common activity related to the certain object with the same role a ; for example, two users take part in the videoconference, two users comment on the same picture, or both of them add the same object to their favourites [25], Figure 2(a). *Object-based relation with different roles* r_{xy}^{ab} , r_{yx}^{ba} is the relation between two users x and y that are connected through the object but their roles a and b towards the object are different; for example, user x comments on a photo (role a , commentator) that was published by y (role b , author) [25], Figure 2(b). The nonzero relation r_{xy}^{ab} entails the nonzero relation r_{yx}^{ba} .

The examples of object-based relations with equal roles are as follows.

- (i) *Commentator-commentator*. This relation is created between user x and user y when both of them have added the opinions about at least one common object; for example, they have commented on the same picture at the photo sharing system or on the same post at the forum.
- (ii) *Favourite-favourite*. Such a relation from user x to y exists if both users have marked at least one common object as their favourite; for example, they have added the same film to their lists of favourites at the multimedia sharing system.

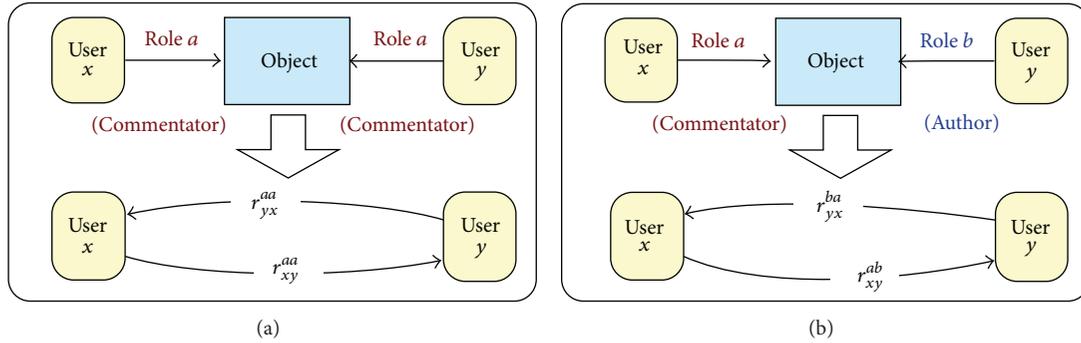


FIGURE 2: The object-based relation with equal roles: *commentator* (a) and different roles: *commentator* and *author* (b).

- (iii) *Author-author*. Such a relation from user x to y exists when they are coauthors of at least one object; for example, they have written a scientific article together.
- (iv) *Membership in the group/forum*. This relation from user x to y exists when both of them belong to at least one group together; for example, they belong to a group that gathers people who like dogs at the photo sharing system.
- (v) *Utilization of keywords to describe objects (tags)*. Such a relation exists between two users if they use at least one common tag to describe their objects; for example, two users are in relation with each other at the photo sharing system when they use a word “cat” to describe some of their photos.

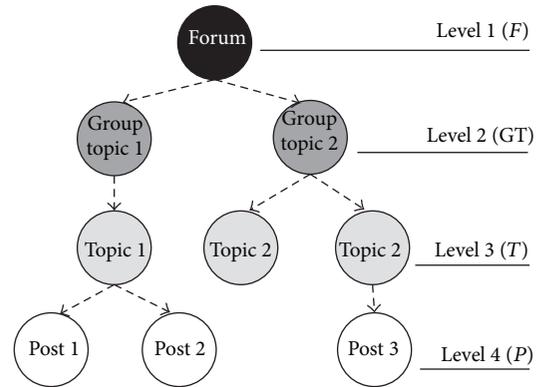


FIGURE 3: Hierarchies between objects in the Internet forum.

On the other hand, the examples of object-based relations with different roles can be as follows.

- (i) *Opinion-author and author-opinion*. These relations between user x and y exist when user x commented on at least one object that is authored by user y .
- (ii) *Favourite-author and author-favourite*. These relations between users x and y exist when user x added to its favourite list at least one object authored by user y .
- (iii) *Citation-author and author-citation*. These relations between users x and y exist when user x quoted at least one object authored by user y .

4. Hierarchies between Objects

In all web-based social networks analysed in the literature, the relationships between users were extracted mainly based on a given type of communication or common activity. For example, if two users send e-mails to each other, then the relationship between them in the social network may be established. However, both user communication and common activities are always related somehow to the objects which serve as a medium in interactions between users and their common activities; see Section 3. This object may be “a message” in the case of e-mail exchange, “a video” in YouTube, or “a topic” in the Internet forums. These objects connect either a pair of users (an e-mail sent to a single recipient)

or many users simultaneously (an e-mail passed to multiple recipients, a video commented on by many users, and a forum with many members). Besides, the IT system may provide many different functions, which can result in various user activities towards objects of different types. For example, an Internet forum may consist of topics aggregated into groups. Topics, in turn, contain a list of posts. Thus, the objects that enable interactions between users are in hierarchical relationships, Figure 3. Depending on the functionalities of the system, users can moderate a topic group, can subscribe to a topic (be a member of the topic), or provide their opinions about posts (play the role of commentator), Figure 10. Hence, users “meet” each other by performing activities towards objects that belong (i) to one specific level in the object hierarchy (many users can comment on a post authored by another user) or (ii) to two different levels of this hierarchy; for example, a moderator of the group topic is in the indirect relation with authors of the posts.

5. Hierarchical Presocial Network

In order to create the hierarchical presocial network (HPSN) based on gathered activity data, first such elements as users, objects, and hierarchy between objects and relations between users and objects need to be extracted. HPSN contains information about relations between users and objects towards

which users performed some activities. The main characteristic of HPSN is that there exist hierarchies between different objects. The whole process of HPSN extraction consists of the following four steps.

- (1) *User extraction.* Users are network nodes both in the presocial network and in the final social network. Users perform different activities towards various types of objects; for example, they send e-mails to each other or comment on the photos uploaded by others. These activities are the basis to create the role of a user in relation to a specific object, for example, author, commentator, and so forth.
- (2) *Object extraction.* Objects are the nodes in both hierarchical and in flat presocial network, that is, elements through which users communicate with each other (e.g., email and phone call) or items towards which users perform some activities (e.g., photo, video, and tag).
- (3) *Extraction of the hierarchy between the objects.* Some objects can be in hierarchical relation with other objects; for example, an object “group of topics” contains one or more “topics” which may include many “posts” (Figure 3). The consequence of the existence of the hierarchy between different objects types is that the objects on the lower level cannot exist without objects on the higher level. These hierarchies exist within HPSN and are removed during the prenetwork flattening process (see Section 6).
- (4) *Extraction of the relations between users and objects.* Relation between a given user a and object X exists if user a performed some activities towards object X ; for example, user a commented on photo X . The type of activity that user performed towards an object is assigned to each relationship.

The concept of HPSN is presented in Figure 4 where the hierarchy between objects has three levels (Z, V, and Y); at each level some objects exist (e.g., at the level Y objects OB Y1 and OB Y2) with which users (a , b , and c) are in different types of relations (Z, V, and Y roles); that is, users performed some activities towards these objects.

6. Flat Presocial Network

The flattening process aims at removing relationships between objects (hierarchies). A consequence of this process is based on the knowledge about existing hierarchies; both new user roles and relationships between users are created. The transformation from the presocial network where the hierarchies between objects exist (HPSN) into the flat presocial network (FPSN) without hierarchies will be performed in the following steps.

- (a) The operator chooses the level in the hierarchy to which the flattening process will be performed—the end level. Note that, after each flattening process, the only object type in FPSN will be the one that is on the

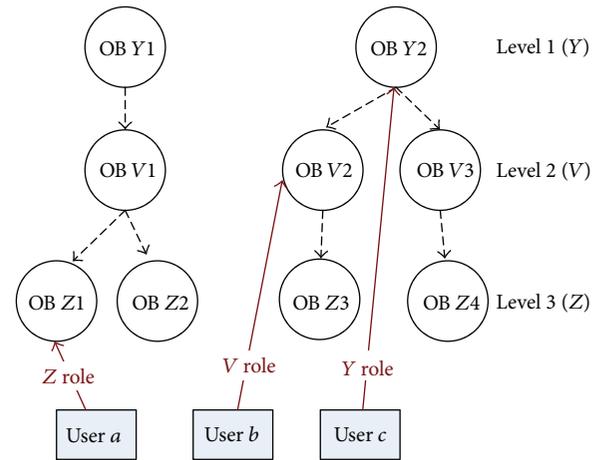


FIGURE 4: The concept of HPSN.

end level selected by the operator and all users will be in relation only towards these objects.

- (b) If there exist levels that are lower in the hierarchy than the end level (Figure 5), then for those levels the *bottom-top* approach is used; that is, we get the following.

- (i) Relationships between people and objects existing on the hierarchy levels that are below the end level (relation user a , OB Z1, and user b , OB V2, in Figure 5) are changed. The relation between a user and an object from the lower level is moved to the upper level by

- (1) identification of an object on the upper level that is “a father” of the object from the lower level (“child”);
- (2) creation of a new relation between the user and “the father” object;
- (3) name of the relation between user and “father” object which is created by adding to the name of the relation user “child” the word that denotes the movement from the lower level. For example, in Figure 5, the relation user a , OB Z1 (“child”), has a name, Z role, and the name of the new relation user a , OB V1 (“father”), is ZV role;
- (4) deletion of the relation between the user and the “child” object from the lower level.

NOTE. This process is repeated for other upper levels until the end level is reached (Figure 5).

- (ii) Relationships between people and objects existing at the end level remain unchanged (relation user c , OB Y2).

The final FPSN presented in Figure 6 is an outcome of the bottom-up approach where the HPSN from Figure 5 is flattened to level 1 (Y).

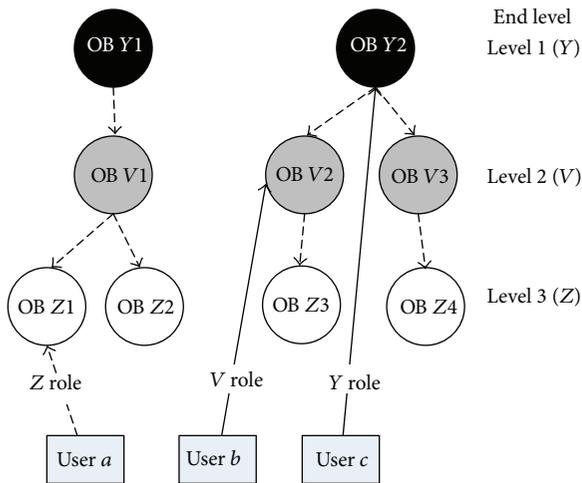


FIGURE 5: Relations between users and objects in the hierarchical presocial network HPSN.

(c) If there exist levels that are upper in the hierarchy than in the end level (Figure 7), then the top-bottom approach is applied for these levels; that is, we get the following.

(i) Relationships between people and objects existing on the hierarchy levels that are above the end level (relation user *b*, OB V2, and user *c*, OB Y2, in Figure 7) are changed. The relation between user and object from the upper level is moved to the lower level by

- (1) identification of all objects on the lower level that are “children” of an object from the upper level (“father”);
- (2) creation of the relation between the user and all “child objects;”
- (3) name of the relation between the user and “child object” which is created by adding to the origin name information about the “child object.” For the example, in Figure 7, in the relation user *b*, OB V2 (“father”), the relation name was V role and the new relation user *b*, OB Z3 (“child”), will have the name VZ role;
- (4) deletion of the relation between the user and “father object” on the upper level.

NOTE. This process is repeated until the end level is reached (Figure 8).

(ii) Relationships between people and objects existing on the end level remain unchanged.

An example of top-bottom approach is presented below where the HPSN from Figure 7 is flattened to level 3 (Z), that is, to FPSN in Figure 8.

The goal of the flattening process is to facilitate the extraction of the unified structure that represents the social connections between pairs of users from user activity data and relations between objects. New types of user roles can

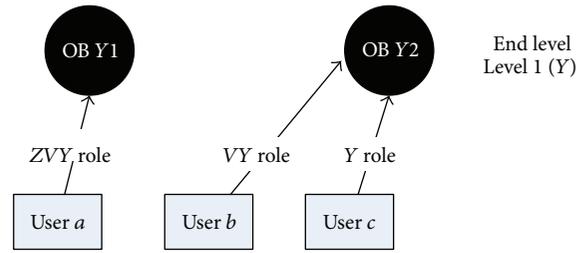


FIGURE 6: Relations between users and objects in the final FPSN after the bottom-top flattening process.

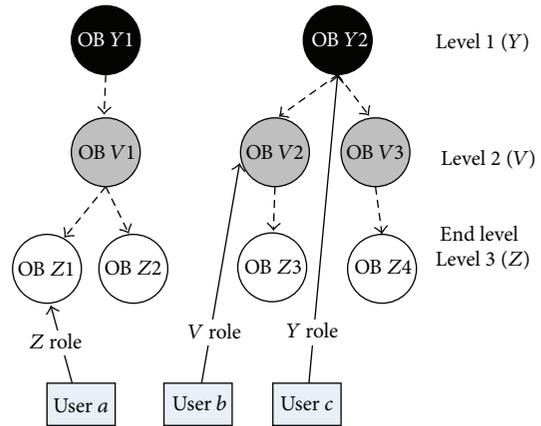


FIGURE 7: Relations between users and objects in the hierarchical presocial network HPSN.

be identified during the flattening process, for example, VZ or YVZ role in Figure 8. The newly obtained knowledge about these roles gives an opportunity to investigate the complex profile of user relationships in more detail and in consequence enables their more comprehensive analysis.

7. Social Network

The flat presocial network structure (FPSN) is used to extract the social network (SN) where the relations user-object from FPSN no longer exist. These connections are converted into direct relations between users in SN. The process consists of the following steps.

- (a) Extraction of SN layers is based on the type of the users’ roles towards objects. Each network layer consists of users and their connections. In a single layer network, there exist object-based relationships of only one type either with equal or different roles (see Section 3).
- (b) The operator chooses which social network layers need to be created.
- (c) The operator chooses the SN model,

n-graph, each layer in the multilayered social network SN is represented by separate social network;

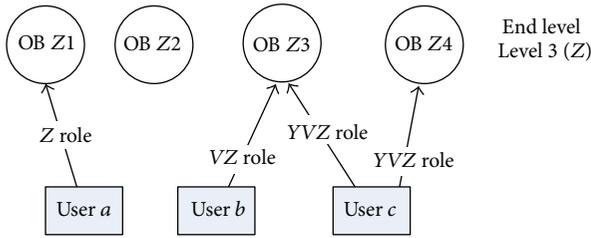


FIGURE 8: Relations between users and objects in the final FPSN after the top-bottom flattening process.

multigraph, all layers are represented by a single social network and different layers are distinguished by different colours of edges (or another labelling mechanism is used).

- (d) Extraction of relations *user_from-user_to* is by calculation of the relationship strengths and colours (labels) between SN nodes (users) using activity data stored in FPSN. There are many possible formulas for calculating the relationship strength. Most of them are based on the normalized quantity of shared user activities towards objects in FPSN (for some of the examples please see [25, 42]).

Social network SN created from FPSN in Figure 8 according to the process described above is presented in Figure 9.

Depending on the goal of analysis, the strength of a relationship can be a static measure calculated based on all available data and taking into consideration the number of activities of a given type. On the other hand, we can take into account time factor and split the data according to the time when activities occurred. In the latter case, the whole period from which the data comes from is divided into time frames and the relationship strength is calculated for each slot separately. The time frame can be created using two approaches as follows.

- (i) *Sliding window*. A user defines the length of the time window (e.g., t_l) and the time interval that is used to move the window (e.g., t_i). In order to extract time frames the whole period of the length t_l is moved by t_i . In consequence the entire dataset is divided into partly overlapping frames. Note that both time window and time interval need to be specified in a way that the period from the start date to end date should be completely covered.
- (ii) *Equal, separate periods*. A user sets the number of periods, for example, k , and then the data is divided into k separate, equal periods according to the dates of activity. This is equivalent to the situation (i) where t_l equals t_i .

After the time windows are created, the weight is assigned to each of them. Usually, the recent periods are more important than the previous ones and because of that greater weight is assigned to those recent time windows.

In this paper we present how to calculate the static version of relationship strength for two different types of relations: (i)

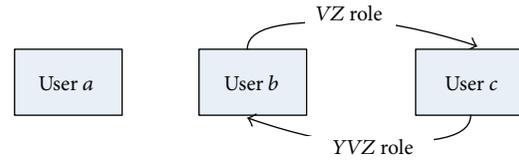


FIGURE 9: Social network created from FPSN presented in Figure 8.

object-based relations with equal roles and (ii) object-based relations with different roles (see Figure 2).

(a) *Object-Based Relationships with Equal Roles*. The object-based relationship with equal roles denotes a connection in which two users are related to each other through the object and their roles towards this object are the same; see Section 3. Note that the same formula is used in order to calculate the connection strength between user x and user y who (i) have commented on at least one common object, (ii) have marked at least one common object as their favourite, (iii) are coauthors of an object, (iv) are in the same group or forum, and (v) have used the same keywords to describe objects. In all of these relations, there is an object on which both users perform specific activity. To calculate the static strength of the relationship, the following formula may be applied:

$$s_{xy}^a = \frac{n_{xy}^a}{n_x^a}, \quad x \neq y, \quad (1)$$

where a is the type of activity that is performed by users towards an object, for example, membership to a group/forum, utilization of a tag to describe objects, coauthorship of an object, commenting on an object, and so forth. n_{xy}^a is the number of common a activities for users x and y performed together, for example, number of groups/forums to which both users x and y belong, the number of tags that both users x and y use commonly or the number of objects that were coauthored by both users x and y , and so forth. n_x^a is the number of a given a activity for user x , for example, the number of groups/forums to which user x belongs, the number of tags used by user x or the number of objects authored by user x , and so forth.

Let us consider the situation in which users of multimedia sharing system utilise tags to describe different multimedia content. In this case an object is a *tag* and relationship between two users is created when they utilise some common tags. Let us assume that the data obtained from the system contains the following information: user x utilised 20 identical tags as user y and user x used 60 tags in total. Then relation strength from user x to y is calculated as follows:

$$s_{ij}^y = \frac{n_{ij}^y}{n_i^y} = \frac{20}{60} = \frac{1}{3}. \quad (2)$$

(b) *Object-Based Relationships with Different Roles*. The object-based relation with different roles denotes a connection in which two users are related to each other through the object and their roles towards this object are different (see

Section 3). For example, one user can comment in a forum in which another user is a moderator and the relationship between users is moderator-commentator. Thus, in the case of calculating the relations strength, we will refer to the relations $activity_type_a-activity_type_b$.

Considering the relation $activity_type_a-activity_type_b$ its strength will be calculated as follows:

$$s_{xy}^{ab} = \frac{n_{xy}^{ab}}{n_x^{ab}}, \quad x \neq y, a \neq b, \quad (3)$$

where a denotes the 1st activity type; b denotes the 2nd activity type (different that a); n_{xy}^{ab} is the number of activities a performed by user x towards objects for which user y performed activity b ; n_x^{ab} is the total number of activities of type a performed by user x towards objects for which any other users performed activity b .

As an example, let us consider the case where a user adds to the list of favourites an object authored by another user. A relationship between two users is created when one user adds to its favourites an object authored by another person. Assume that the following data is available in the system: user x added to favourites 20 objects authored by user y . User x added to favourites 60 objects in total. The objects of user y were added to favourites by other 30 times in total. Moreover, a means the activity “authored by” and b means the activity “added to favourite by”.

The relation strengths are calculated as follows:

$$\text{from user } x \text{ to } y: s_{xy}^{ba} = n_{xy}^{ba}/n_x^{ba} = 20/60 = 1/3;$$

$$\text{from user } y \text{ to } x: s_{yx}^{ab} = n_{yx}^{ab}/n_y^{ab} = 20/30 = 2/3.$$

8. Example of Flattening Process

One of the examples where hierarchies between objects exist is the Internet forum where people can create their own topics that contain posts added by users. The hierarchy between objects within a forum and the activities that can be performed towards these objects are presented in Figure 10.

The hierarchy that will be used for our case study is forum-topic-post. Both relationships between objects and between users and objects in the exemplary hierarchical prenetwork are presented in Figure 11. In order to create the flat presocial network (Figure 12) we perform the flattening process after which the relationships between objects will be removed but at the same time on other levels new relationships between users and objects will be created.

Two types of flattening process can be considered: bottom-top and top-bottom (see Section 6). We present here the bottom-top flattening process in which the relationships will be moved to the highest forum level; that is, the forum will be the final level.

The bottom-top approach applied to HPSN from Figure 11 results in several new relationships and roles (Figure 12). Some examples are enumerated as follows.

- (i) User A is an author of post 1.2.1 in the topic 1.2 at the forum 1 and also of post 2.1.1 in the topic 2.1 at

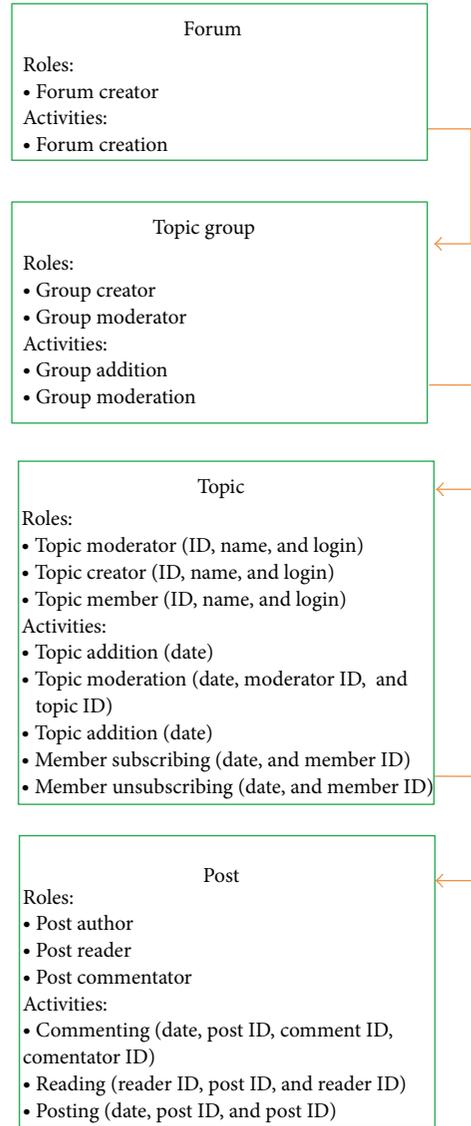


FIGURE 10: The hierarchy of objects in the forum.

forum 2. Then a new relation between user A and forum is created: $PTF \text{ Is Author } (PT, PostTopicForum)$, a new name of relation; see Section 6); that is, user A authored at least one post in a topic that is in a given forum. A similar approach is applied to user B , who is an author of post 1.1.1 in the topic 1.1 and to user C , an author of post 2.2.2 in the topic 2.2. Moreover, user D is an author of posts 1.1.2 and 2.2.1. This is flattening (movement) of authorship activities on posts (role $P \text{ Is Author}$) to the top level forum.

- (ii) User B is a commentator of post 2.1.1. A new relation between user B and the forum is created: $PTF \text{ Is Commentator}$. The same is done for user D , a commentator of posts 1.2.1, 2.1.1, and 2.2.2. In this way, the role $P \text{ Is Commentator}$ is moved to the forum level.
- (iii) User A is a moderator of topic 1.1 at the forum; then a new relation between user A and the forum is created:

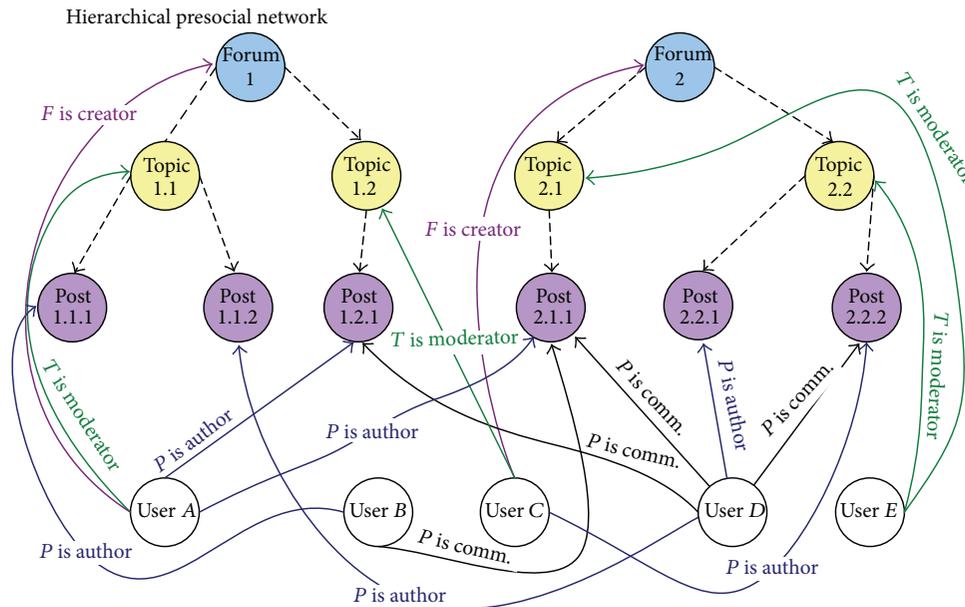


FIGURE 11: Relations between users and objects together with users' roles towards the objects.

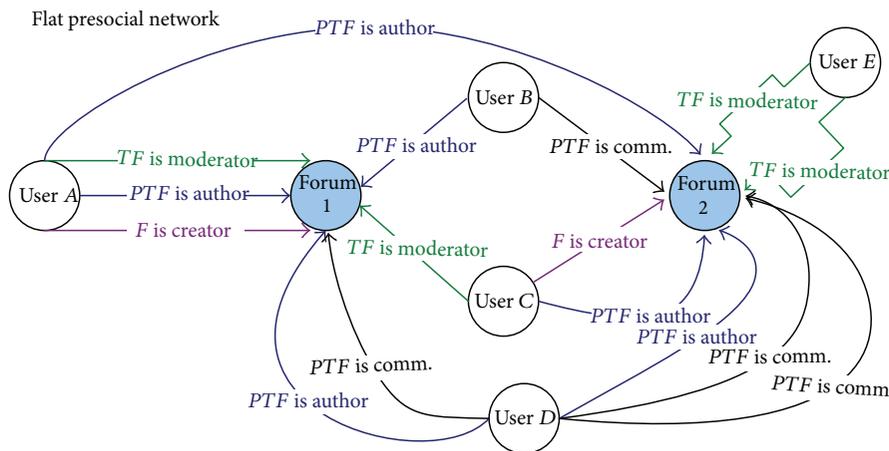


FIGURE 12: Flat presocial network (FPSN): the result of flattening process applied to HPSN from Figure 11.

TF Is Moderator (TF, TopicForum, a new name of relation; see Section 6). The similar method is utilized for user C, the moderator of topic 1.2, and user E, the moderator of topics 2.1 and 2.2. This is flattening of roles from the topic level to the final forum level.

- (iv) *User A* is a creator of the forum 1 and *user C* is a creator of forum 2; then the existing relationship stays unchanged: *F Is Creator*.

Two different layers of the final multilayered SN derived from the flat presocial network FPSN (Figure 12) are presented in Figures 13 and 14. For instance, the relation between two users exists if one user was the moderator of the post that was commented on by another user, Figure 13; for example, *user A* moderates the post commented on by *user D* so there is a relation *moderator-commentator* from *A* to *D* in the final social network SN. This relation is an object-based

relationship with different roles. In Figure 14, another layer *PTF Is Author, PTF Is Author*, is presented. This is a layer in which the extracted relationships are object-based with equal roles. The relationships' strengths presented on both figures are calculated using formulas from Section 7.

Note that *user D* and *user E* are connected only because of flattening process (Figure 13). The same is with *user E* and *user B* (Figure 13) and *user A* and *user C* (Figure 14). Any of these relationships would be revealed if the flattening process did not take place.

9. Experiments

The real-world dataset used for experiments was obtained from the social web site *extradom.pl*. The analysed dataset covers the period from August 21, 2008, to January 8, 2010.

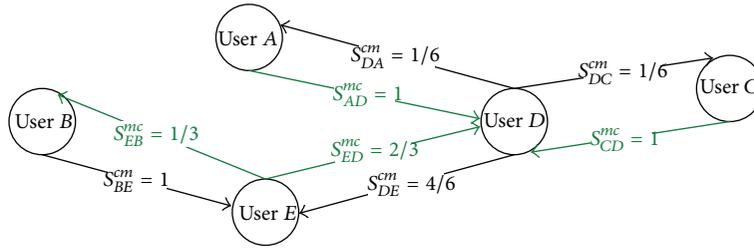


FIGURE 13: Example of the *PTF Is Commentator* and *MF Is Moderator* layer of the final SN extracted from the flat presocial network FPSN (Figure 12).

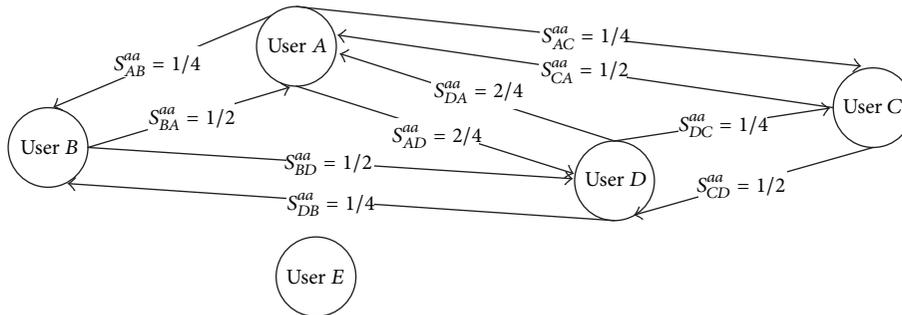


FIGURE 14: Example the *PTF Author* *PTF Author* layer of the final SN extracted from the flat presocial network FPSN (Figure 12).

Before the hierarchical presocial network HPSN was created, the dataset was cleansed and validated. Several rules were applied in the cleansing phase. Two most important ones were as follows:

- (i) each object must have creation date;
- (ii) each object must be assigned to its creator.

There were 104,625 users registered in the portal, but only 4.25% (4,404) of them were active on forum. The number of different types of objects in the forum is shown in Table 1. One or more activity types were identified for each object type. Objects with activities that were performed towards them are shown in Table 2.

As we can see in Table 2, there are two activity types (*topic group moderation* and *post reading*) that are not present in the dataset used for experiments and thus are not included in the further analyses. Additional assumptions that have been made and which helped to detect some of activities are as follows:

- (i) a user who creates the first post in the topic will be treated as a topic creator;
- (ii) creation of the first topic in the group is simultaneously treated as the creation of the entire group;
- (iii) users who create their first post in the topic will be automatically subscribed to this topic.

Table 3 summarizes the profile of the hierarchical presocial network (HPSN) that was created from the extradom.pl dataset.

Three distinct flattening processes (see Section 6) with three separate final object levels, topics group, topic, and

TABLE 1: Objects quantity in the experimental dataset.

Object type	Number of objects
Forum	1
Topic group	692
Topic	2,336
Post	13,272
Comment	49

TABLE 2: Activities assigned to object types.

Object type	Activity type	Number of activities
Forum	Forum creation activity	1
Topics group	Topic group addition	692
	Topic group moderation	0
Topic	Topic addition	2,336
	Topic member subscribing	5,788
Post	Post reading	0
	Post authoring	13,272
Comment	Post commenting	49

post, have been applied to the hierarchical presocial network HPSN. Some of the activities were multiplied after the flattening process (for detailed statistics please see Table 4). Such situation takes place when the hierarchical presocial network is flattened to the object type, which is not the highest level in the hierarchy. For example, when HPSN was flattened to the topic groups (level 2 in the hierarchy, see Figure 3), the forum creation activity was multiplied by 692 because there were distinct 692 groups in the dataset.

TABLE 3: Activities in the hierarchical presocial network.

Activity type	Number of activities	Number of users with the activity	Percentage of total users with a given activity
Forum creation activity	1	1	0.02%
Topic group addition	692	266	6.04%
Topic addition	2,336	1,464	33.24%
Topic member subscribing	5,788	4,359	98.98%
Post posting	13,272	4,359	98.98%
Post commenting	49	39	0.89%

TABLE 4: Number of activities before and after flattening; “+” denotes activities which were created during a given flattening process.

Activity type	Number of activities before flattening (in HPSN)		Number of activities after flattening (in FPSN)				
		New	Final objects: topic groups	New	Final objects: topics	New	Final objects: posts
Forum creation activity	1	+	692	+	2336	+	1,3272
Topic group addition	692		692	+	2336	+	13,272
Topic addition	2,336		2,336		2336	+	13,272
Topic member subscribing	5,788		5,788		5788		793,245
Post authoring	13,272		13,272		13,272		13,272
Post commenting	49		49		49		49

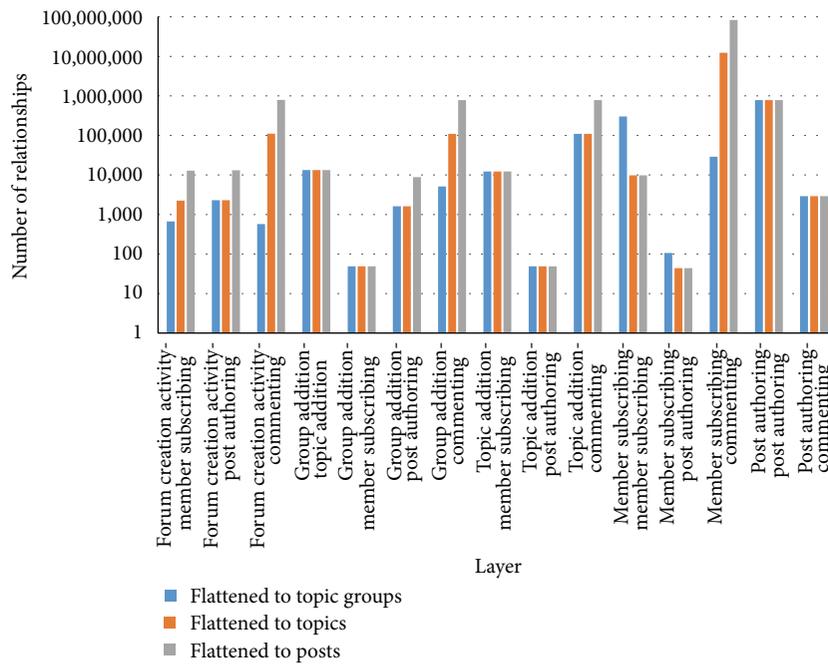


FIGURE 15: Number of activities (relationships) within created layers for three different flattening processes.

Once the flattening process has been accomplished, the separate layers in the multilayered social network were identified and relationships between users within these layers were extracted (see Section 7). Both layers and number of distinct relationships existing within each layer are presented in Table 5.

As a result of the flattening process 14 new layers in the multilayered social network were created. Additionally,

percentages of new user relationships in SN are 90% in the case of topic groups as the final object, 10% for topics, and 3% for posts (see Table 5 and Figure 15). Note that these new relationships would not be visible without the flattening process. It means that the method of preprocessing with flattening of object relations reveals completely new knowledge about the complexity of connections between people.

TABLE 5: Layers in SN and their profile for three different flattening processes.

Layer	Number of pairs of users (relationships) with common activities on the layer of SN									
	New	Moved	Topic groups	Moved	Topics	Topics/ topic groups	Moved	Posts	Posts/ topics	
Forum creation activity—group addition	+		671		2,252	3.36		12,922	5.74	
Forum creation activity—topic addition	+		2,313		2,313	1.00		13,157	5.69	
Forum creation activity—member subscribing	+		5,760		111,544	19.37		792,732	7.11	
Forum creation activity—post authoring	+		13,233		13,233	1.00		13,233	1.00	
Forum creation activity—commenting	+		49		49	1.00		49	1.00	
Group addition—topic addition	+		1,614		1,614	1.00		8,821	5.47	
Group addition—member subscribing	+		5,116		109,377	21.38		780,312	7.13	
Group addition—post authoring	+		12,206		12,206	1.00		12,206	1.00	
Group addition—commenting	+		49		49	1.00		49	1.00	
Topic addition—member subscribing		+	109,376		109,376	1.00	+	780,445	7.14	
Topic addition—post authoring	+		302,961		9,849	0.03		9,849	1.00	
Topic addition—commenting	+		1,060		44	0.04		44	1.00	
Member subscribing—member subscribing		+	289,740	+	12,342,690	42.60	+	83,495,278	6.76	
Member subscribing—post authoring	+		779,973		779,973	1.00		779,973	1.00	
Member subscribing—commenting	+		2,903		2,903	1.00		2,903	1.00	
Post authoring—post authoring	+		2,449,226		376,978	0.15		0	0	
Post authoring—commenting		+	9,340	+	1,703	0.18		44	0.03	
			3,985,590		13,876,153	3.48		86,702,017	6.25	
Sum:			New: 90%		New: 10%			New: 3%		
			Moved: 10%		Moved: 90%			Moved: 97%		

10. Conclusions

The wide variety and availability of Web 2.0 systems, where users can interact with each other and perform different types of activities, give us an opportunity, by analysing the large-scale data gathered in these systems, to better understand human social behaviour. A very interesting research problem is to investigate social connections that emerge between people based on their shared activities. However, the extraction of these relations is not a trivial task. The main reason is that user behaviour in such systems is often very complex due to the variety of available services and functionalities. As presented in the paper, people can perform different activities towards different objects. Additional challenge is that the relationships existing between these objects can form a hierarchical structure. In this paper, we propose the process to extract the multilayered social network from the data about both user behaviours and relations between objects. The whole method consists of three main phases: (i) extraction of the hierarchical presocial network HPSN, (ii) creation of the flat presocial network PFSN, and finally (iii) creation of the multilayered social network SN. We believe that such systematic approach to the problems is necessary to be able to cope with the massive volume of data being generated by social-based systems every day. Moreover, the proposed process is generic and robust in a way that it is able to accommodate new ways of interactions between users.

The new flattening concept enables discovering in the multilayered social network new layers with new types of relationships which otherwise would not be available for analysis. It is possible due to the presented above process in which the object hierarchy is removed. Thus, the new method of preprocessing enables revealing new information, which is invisible in the regular analysis of user activities and this in turn opens new possibilities for network analysis. The experiments confirmed that some new types of relations between users can be extracted in the flattening process. This enables a deeper insight into analysis of multilayered social networks as more information is included in the final network structure.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

The authors are indebted to Elżbieta Kukła, Ph.D. holder, and Tomasz Filipowski, M.S. holder, for their valuable discussion. The work was partially supported by Fellowship cofinanced by the European Union within European Social

Fund, The European Commission under the 7th Framework Programme, Coordination, and Support Action, Grant Agreement no. 316097, ENGINE, European Research Centre of Network intelligence for INnovation Enhancement <http://engine.pwr.wroc.pl/>, and The National Science Centre, Decision no. DEC-2013/09/B/ST6/02317.

References

- [1] A. Barrat, M. Barthelemy, and A. Vespignani, *Dynamical Processes on Complex Networks*, Cambridge University Press, Cambridge, UK, 2008.
- [2] L. A. Adamic and E. Adar, "Friends and neighbors on the Web," *Social Networks*, vol. 25, no. 3, pp. 211–230, 2003.
- [3] J. Scott, *Social Network Analysis: A Handbook*, SAGE Publications, London, UK, 2000.
- [4] S. H. Strogatz, "Exploring complex networks," *Nature*, vol. 410, no. 6825, pp. 268–276, 2001.
- [5] S. Wasserman and K. Faust, *Social Network Analysis: Methods and Applications*, Cambridge University Press, New York, NY, USA, 1994.
- [6] D. J. Watts and S. H. Strogatz, "Collective dynamics of "small-world" networks," *Nature*, vol. 393, no. 6684, pp. 440–442, 1998.
- [7] R. Hanneman and M. Riddle, *Introduction to Social Network Methods*, University of California, Riverside, Calif, USA, 2005, <http://faculty.ucr.edu/~hanneman/nettext/>.
- [8] M. S. Baptista, C. Grebogi, and R. Köberle, "Dynamically multilayered visual system of the multifractal fly," *Physical Review Letters*, vol. 97, no. 17, Article ID 178102, 2006.
- [9] P. Bródka, P. Kazienko, K. Musiał, and K. Skibicki, "Analysis of neighbourhoods in multi-layered dynamic social networks," *International Journal of Computational Intelligence Systems*, vol. 5, no. 3, pp. 582–596, 2012.
- [10] P. Bródka, K. Skibicki, P. Kazienko, and K. Musiał, "A degree centrality in multi-layered social network," in *Proceedings of the International Conference on Computational Aspects of Social Networks (CASoN '11)*, pp. 237–242, October 2011.
- [11] P. Bródka, P. Stawiak, and P. Kazienko, "Shortest path discovery in the multi-layered social network," in *Proceedings of the International Conference on Advances in Social Networks Analysis and Mining (ASONAM '11)*, pp. 497–501, July 2011.
- [12] M. De Domenico, A. Sole-Ribalta, E. Cozzo et al., "Mathematical formulation of multi-layer networks," *Physical Review X*, vol. 3, Article ID 041022, 2013.
- [13] P. Kazienko, P. Brodka, and K. Musial, "Individual neighbourhood exploration in complex multi-layered social network," in *Proceedings of the 3rd IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT '10)*, pp. 5–8, September 2010.
- [14] M. C. Pasqual and O. L. de Weck, "Multilayer network model for analysis and management of change propagation," *Research in Engineering Design*, vol. 23, no. 4, pp. 305–328, 2012.
- [15] F. Darabi Sahneh, C. Scoglio, and P. Van Mieghem, "Generalized epidemic mean-field model for spreading processes over multilayer complex networks," *IEEE/ACM Transactions on Networking*, vol. 21, no. 5, pp. 1609–1620, 2013.
- [16] M. Berlingerio, M. Coscia, F. Giannotti, A. Monreale, and D. Pedreschi, "The pursuit of hubbiness: analysis of hubs in large multidimensional networks," *Journal of Computational Science*, vol. 2, no. 3, pp. 223–237, 2011.
- [17] M. Coscia, F. Giannotti, and D. Pedreschi, "A classification for community discovery methods in complex networks," *Statistical Analysis and Data Mining*, vol. 4, no. 5, pp. 512–546, 2011.
- [18] M. Zignani, C. Quadri, S. Gaitto, and G. P. Rossi, "Exploiting all phone media? A multidimensional network analysis of phone users' sociality," <http://arxiv.org/abs/1401.3126>.
- [19] I. Abraham, S. Chechik, D. Kempe, and A. Slivkins, "Low-distortion inference of latent similarities from a multiplex social network," in *Proceedings of the 24th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA '13)*, pp. 1853–1872, January 2013.
- [20] N. S. Contractor, P. R. Monge, and P. M. Leonardi, "Multi-dimensional networks and the dynamics of sociomateriality: bringing technology inside the network," *International Journal of Communication*, vol. 5, no. 1, pp. 682–720, 2011.
- [21] S. Gómez, A. Díaz-Guilera, J. Gómez-Gardeñes, C. J. Pérez-Vicente, Y. Moreno, and A. Arenas, "Diffusion dynamics on multiplex networks," *Physical Review Letters*, vol. 110, no. 2, Article ID 028701, 2013.
- [22] S. Lee and P. Monge, "The coevolution of multiplex communication networks in organizational communities," *Journal of Communication*, vol. 61, no. 4, pp. 758–779, 2011.
- [23] M. Girvan and M. E. J. Newman, "Community structure in social and biological networks," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 99, no. 12, pp. 7821–7826, 2002.
- [24] N. Agarwal, M. Galan, H. Liu, and S. Subramanya, "WisColl: collective wisdom based blog clustering," *Information Sciences*, vol. 180, no. 1, pp. 39–61, 2010.
- [25] P. Kazienko, K. Musiał, and T. Kajdanowicz, "Multidimensional social network in the social recommender system," *IEEE Transactions on Systems, Man, and Cybernetics A: Systems and Humans*, vol. 41, no. 4, pp. 746–759, 2011.
- [26] J. R. Tyler, D. M. Wilkinson, and B. A. Huberman, "Email as spectroscopy: automated discovery of community structure within organizations," in *Communities and Technologies*, pp. 81–96, Kluwer BV, Deventer, The Netherlands, 2003.
- [27] V. D. Blondel, J.-L. Guillaume, R. Lambiotte, and E. Lefebvre, "Fast unfolding of communities in large networks," *Journal of Statistical Mechanics: Theory and Experiment*, vol. 2008, no. 10, Article ID P10008, 2008.
- [28] P. Kazienko, D. Ruta, and P. Bródka, "The impact of customer churn on social value dynamics," *International Journal of Virtual Communities and Social Networking*, vol. 1, no. 3, pp. 60–72, 2009.
- [29] B. A. Huberman, D. M. Romero, and F. Wu, "Social networks that matter Twitter under the microscope," *First Monday*, vol. 14, no. 1, 2009.
- [30] N. B. Ellison, C. Steinfield, and C. Lampe, "The benefits of facebook "friends": social capital and college students' use of online social network sites," *Journal of Computer-Mediated Communication*, vol. 12, no. 4, pp. 1143–1168, 2007.
- [31] A. L. Traud, E. D. Kelsic, P. J. Mucha, and M. A. Porter, "Comparing community structure to characteristics in online collegiate social networks," *SIAM Review*, vol. 53, no. 3, pp. 526–543, 2011.
- [32] X. Cheng, C. Dale, and J. Liu, "Statistics and social network of YouTube videos," in *Proceedings of the 16th International Workshop on Quality of Service (IWQoS '08)*, pp. 229–238, June 2008.

- [33] A. Capocci, V. D. P. Servedio, F. Colaiori et al., “Preferential attachment in the growth of social networks: the internet encyclopedia Wikipedia,” *Physical Review E: Statistical, Nonlinear, and Soft Matter Physics*, vol. 74, no. 3, Article ID 036116, 2006.
- [34] P. Mika, “Flink: semantic Web technology for the extraction and analysis of social networks,” *Web Semantics*, vol. 3, no. 2-3, pp. 211–223, 2005.
- [35] Y. Matsuo, J. Mori, M. Hamasaki et al., “POLYPHONET: an advanced social network extraction system from the Web,” *Web Semantics*, vol. 5, no. 4, pp. 262–278, 2007.
- [36] P. Bródka and P. Kazienko, *Multi-Layered Social Networks. Encyclopedia of Social Network Analysis and Mining*, Springer, 2014.
- [37] P. Kazienko, P. Bródka, K. Musiał, and J. Gaworecki, “Multi-layered social network creation based on bibliographic data,” in *Proceedings of the 2nd IEEE International Conference on Social Computing (SocialCom '10)*, pp. 407–412, August 2010.
- [38] R. Michalski, P. Kazienko, and J. Jankowski, “Convince a dozen more and succeed—the influence in multi-layered social networks,” in *Proceedings of the 9th International Conference on Signal Image Technology & Internet based Systems*, pp. 499–505, IEEE Computer Society, 2013.
- [39] T. J. Zagenczyk, R. L. Purvis, M. K. Shoss, K. L. Scott, and K. S. Cruz, “Social influence and leader perceptions: multiplex social network ties and similarity in leader-member exchange,” *Journal of Business and Psychology*, 2013.
- [40] A. Lomi, D. Lusher, P. E. Pattison, and G. Robins, *Inter-Organizational Hierarchies, Social Networks, and Identities in Multi-Unit Organizations*, American Sociological Association, New York, NY, USA, 2007, http://www.allacademic.com/meta/p183413_index.html.
- [41] K. Musiał and P. Kazienko, “Social networks on the Internet,” *World Wide Web*, vol. 16, no. 1, pp. 31–72, 2013.
- [42] G. Caldarelli and A. Vespignani, Eds., *Large Scale Structure and Dynamics of Complex Networks, from Information Technology to Finance and Natural Science, Complex Systems and Interdisciplinary Science*, vol. 2, World Scientific, 2007.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

