



**HAL**  
open science

## Achieving and maintaining important roles in social media

Anna Zygmunt, Jaroslaw Koźlak, Bogdan Gliwa, Maria Stojkow, Dorota Żuchowska-skiba, Yves Demazeau

► **To cite this version:**

Anna Zygmunt, Jaroslaw Koźlak, Bogdan Gliwa, Maria Stojkow, Dorota Żuchowska-skiba, et al.. Achieving and maintaining important roles in social media. *Information Processing and Management*, 2020, 57 (3), pp.102223. 10.1016/j.ipm.2020.102223 . hal-02490009

**HAL Id: hal-02490009**

**<https://hal.science/hal-02490009>**

Submitted on 8 Dec 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Achieving and maintaining important roles in social media

Anna Zygmunt, Jarosław Koźlak, Bogdan Gliwa

*AGH University of Science and Technology  
Faculty of Computer Science, Electronics and Telecommunications  
Al. Mickiewicza 30, 30-059 Kraków, Poland  
{azygmunt,kozlak}@agh.edu.pl, bgliwa@gmail.com*

Maria Stojkow, Dorota Żuchowska–Skiba

*AGH University of Science and Technology  
Faculty of Humanities  
Al. Mickiewicza 30, 30-059 Kraków, Poland  
{stojkoff,zuchowskadorota}@gmail.com*

Yves Demazeau

*CNRS – LIG,  
700 avenue Centrale  
Domaine Universitaire – CS 40700  
38058 Grenoble CX 9  
Grenoble, France  
Yves.Demazeau@imag.fr*

---

## Abstract

A significant problem common for different domains of applications is an issue of obtaining and keeping an influential position in the respective social media society. In this paper a new approach is proposed based on the analysis of roles of users in groups identified within society. Three different dimensions of user behavior are considered as key elements of these roles: their activity, influence and cooperativeness/competition. Taking into account measures describing these dimensions, a set of roles characterizing behaviors of users in groups is formulated.

We propose an original set of roles with their justification in sociological models, develop an easy extendable model of a social system and conduct experiments to allow us to define patterns describing stability and variability of given roles as well as statistics of transitions in time between these considered roles. To define the roles, we took into account different features of user interactions, both quantitative and qualitative. We propose an integrated approach to the analysis of role changes in the context of group evolution. We consider behavior of users in groups with different sizes and differences between them. We also analyse the stability of individual roles players by users in groups and often occurring transitions between individual roles. The obtained results, interpreted also from the sociological point of view, allow the formulation of general recommendations on which behaviors of users could ensure achieving and maintaining influential roles in social media. The most frequent patterns of transitions between roles are identified and significant similarities between them for two considered blog portals are described. The approaches and methods of analysis presented in the paper may be applied to support decisions leading to obtaining and maintaining influential positions in social media, which may be useful for the promotion of goods and services, leading business or political campaigns.

*Keywords:* social roles, role transitions, social network dynamics, social groups, social network, social media

---

## 1. Introduction

Nowadays, different forms of social media (social networking, blogs, forums, microblogging, bookmarking sites, social news, media sharing) have become an important source of knowledge regarding human behavior and processes

5 taking place among their users (Agarwal and Liu (2009), Benevenuto et al. (2009), Kumar et al. (2013), Zafarani et al. (2014), Teh et al. (2015), Chang et al. (2017), Hromic and Hayes (2018)). Research on the activities of individuals in social networks, including social media, allow the learning of the mechanisms that influence the organization of modern societies by identifying actors playing an important role in maintaining the stability of patterns of information flow in social networks. As a result, it makes it possible to discover the impact of individual interaction patterns on society organization (Makagon et al. (2012)). It is important to understand rules of behavior of the social portal which are consequences of actions performed by a population of users on the portal. What we are going to be particularly interested in is the identification of influential roles that exist within the realm of social media. This requires determining users with social impact.

10 In recent years, the dynamic development of social media has increased the interest of researchers in the subject of the social impact of individuals on the immediate environment. The analyses carried out in this trend refer to the identification of influential users in online groups (Kempe et al. (2003); Kim and Han (2009); Kimura et al. (2010)). This is important because influential individuals can play key roles in disseminating information, thereby influencing the opinions and views of other group members (Eirinaki et al. (2012)).

15 Factors affecting individuals in virtual social networks are very similar to those in real life and have existed before the emergence of the Web 2.0 technology that enables communities to form in the Internet space (Libai et al. (2010); Eccleston and Griseri (2008)). Thus, this allows referring similar analytical categories to the identification of influential individuals in both real and virtual world. In the real world, we consult different people in terms of their views and opinions on various topics. In the virtual, such a role is played by community sites; we use advice and recommendations posted on blogs when buying brand-specific products, selecting a movie, and seeking information or advice on current fashion trends. Users whose posts are known and frequently chosen by network users can be described as influential bloggers (Agarwal et al. (2012)). They play a key role in disseminating information (Varlamis et al. (2013)). They can exert a significant influence on the course of political, social and marketing campaigns by spreading their opinions, views and comments supporting new products or ideas (Farrell and Drezner (2007); Kempe et al. (2003)).

20 Users interact with one another locally, and as a result of these activities, groups are formed. In many publications the groups may also be called “communities”, in this paper we decided to use the term “groups”. Each user may belong to numerous such groups, behave differently in each of them and play different roles. The relations between users in social media may significantly change over time, so it is important to consider this in the analysis. These changes may also cause the reorganisation of existing groups in society. An active or influential member of a group in a given time period may later lose their significant role as a result of various conditions. It might be useful to discover the factors that decide whether a given user will keep their important role during a longer period of time.

### 30 1.1. Research goal and main contribution

The goal of the work is to understand the rules of behavior of the influential users in the social media systems in the context of the evolution of their behavior and the stability of the identified features of the users. We describe the behavior of the users in groups, to which they belong to, using a set of roles which are assigned to them. We identify how the influential users differ among themselves and what is an impact of these differences on the stability of their behaviors and rules describing their changes.

45 An important original contribution of this work is a holistic view of the problem of roles performed by users in groups and the evolution of these roles over time. There are other papers on these issues, which, however, deal with the evolution of groups and the evolution of roles separately. To our best knowledge such detailed analyses of the evolution of user roles in groups, and in particular the frequency of transitions between roles in time had not been made so far. The novelty of the research presented in the article is to show how the factors such as influence, cooperation and user activity that affect the roles and positions of the individual in groups are used together to identify roles of users and to show how they change over time.

50 We assume that to understand the rules of user role evolution in groups it is important to identify the most common transitions which occur between roles over time. We believe that so far this problem has not been so thoroughly and comprehensively analyzed. In this paper the most frequent patterns of transitions between roles on two different social media portals are identified and significant similarities between both these cases are described. We also identified which roles are more frequently maintained by the users and which change easier to other roles. These observations

55 may have a significant practical meaning while trying to prepare politics for the user how they should behave to gain significant importance or a position with given features in groups in social media. Meanwhile, this is related to a significant problem how the user in social media should behave to obtain or maintain an influential position in the group.

Our original contributions consists of the following, more detailed elements:

- 60 • We propose *an original set of roles with their justification in sociological models* allowing us to differentiate users considering their especially important characteristic dimensions (*levels of activity, influence and cooperativeness/competition*), we decided to base the representation of roles on these three dimensions assuming that they are especially important for the classification of the most essential characteristics of user behavior in social media.
- 65 • We develop *an easy extendable model of a social system*, considering different roles assigned to users and information about transitions between roles, specify precisely described elements of the system and links between them, a development of such a model facilitates choice of the experimental scenarios and their interpretation, identification of considered states of the system and dependencies between elements which need to be analyzed.
- 70 • We conduct experiments to allow us to *define patterns describing stability and variability of given roles* – chances of obtaining the given role or chances of keeping or obtaining an influential role. We then can draw conclusions about which behaviors give a significant chance of promotion and which give a high level of safety/security to maintain their influential roles.
- 75 • We find *temporal patterns of frequent transitions* between roles of users on the social portal. It helps to elaborate on the right strategies that allow the achieving of sufficiently significant influence in society, which may be used later in different ways, such as business (promotion of products or companies) or political marketing (promotion of politics or political ideas).

The study of social roles and their interrelationships as well as the factors affecting their positions in the Internet space is of great importance for the development of social sciences. By conducting analyses deriving from the classical sociological concepts of social roles, it is possible to get to know the structure of online communities (Agrawal et al. (2003)). Previous studies of social roles adopted by actors in online communities show that there is a relationship 80 between the roles of individuals and the structure of the whole group (Agrawal et al. (2003), Herrmann et al. (2004)). This is due to the fact that the analysis of social roles makes it possible to detect behaviors based on cooperation or competition in networks (Golder and Donath (2004)). It also allows the discovery of mechanisms of internal dynamics in online communities and factors affecting changes in position in online groups (Rullani and Haefliger (2013)). As a result the research that is based on the concept of social roles show mechanisms that contribute to a 85 better understanding of the functioning of online communities and groups and the processes that take place within them that affect the formation of their structure (Benamar et al. (2017)). The analyses presented in the article are part of the research on the significance of the roles of actors in the blogosphere in the context of their activities. Thanks to this, they contribute to a better understanding of the processes related to the functioning of the social space of the Internet, because they enable to learn about factors that allow an individual to perform influential roles in the online 90 community.

The remainder of this article is organized as follows. In Section 2, we introduce areas related to the analysis of user behavior in groups identified in social media . Section 3 explains our methodology of social media analysis with the identification of groups and roles of users, taking into consideration their evolution of time. Section 4 covers concepts of role dimensions and role description. In Section 5, a social system model is presented. The analyze of obtained 95 results is described in Section 6, where we also discuss the datasets and methodology of experiments. Finally, we conclude the work and discuss open issues and future directions in Section 7.

## 2. Related work

We start this chapter with the description of the novelty of our approach and differences between the method presented in this paper and the works published so far (Section 2.1).

100 Subsequent sections contains the research domain overview and cover selected problems of social network analysis (such as the specific features of the blogosphere and the problem of identification of groups - Section 2.2) and after focuses on the discussions of the notions strongly related to the subject of our work, such as influence (Section 2.3), activity (Section 2.4) and cooperativeness (Section 2.5), which are used after to develop definitions of roles introduced in Section 4, the overview is finished by the section with the discussion of social roles (Section 2.6).

### 105 2.1. Novelty of our work in the context of existing research

We believe that in order to understand user behaviors on social networks, due to their large size and very diverse nature of interactions among users, e.g. discussion topics in the case of blog portals, an analysis of roles at the level of the whole community has significant limitations and does not sufficiently reflect actual tasks/behaviors of the user. Analyzing user roles at the level of individual groups may be more useful for such purposes. We are interested in 110 determining how stable these behaviors are in the groups and describing frequent patterns that can be followed by changes in these behaviors. We have acknowledged that it is useful to consider three dimensions of user behaviors to identify roles: influence, activity, and willingness to cooperate understood as the engagement in commenting on the content shared by other users, not only on one's own content. While reviewing the research areas which fall within the scope of our work, we did not notice any papers that would focus on the problem defined in this way. While 115 there are numerous works that identify groups, there are very few that study transitions of roles in groups in time, and those which do so are limited to an analysis of stability only (Rossi et al. (2013)). The authors raise their own issue of the significance of such research (Pei et al. (2019)). However, these works use a completely different way of identifying roles from the one we have adopted (Rossi et al. (2013), Abnar et al. (2014), Rossi and Ahmed (2014), Arockiasamy et al. (2016), Pei et al. (2018)), as they perform an analysis of similar neighborhood configurations based 120 on association matrices. Although this approach has some advantages, it also has important limitations related to the difficulty of providing explanations of the role definitions, except for very classic cases (a star, a bridge, etc.) (Pei et al. (2018)). In our approach, roles are precisely defined and interpreted in relation to these three basic dimensions considering their significance in the blogosphere community. Moreover, to our knowledge there are no papers (except (Abnar et al. (2014))) in which roles are analysed taking into account the specificity of the given social medium. 125 We carry out an analysis of transitions between all the roles under consideration. We found only one paper on the joint analysis of the roles and community (Pei et al. (2019)), but a different role identification method based on the neighborhood matrix is used there. The paper does not contain any analysis of role dynamics, role transitions or their stability.

To define the roles, we took into account the much greater complexity of the interaction, both quantitative and 130 qualitative then the work presented in the papers described in previous sections. We propose a much more integrated approach to the analysis of role changes in the context of group evolution. We consider behavior of users in groups with different sizes and differences between them. We also analyse the stability of individual roles players by users in groups and often occurring transitions between individual roles. Understanding the mechanisms of influence on position in social networks plays an important role in social analysis and allows us to understand the actions taken by 135 social actors online. The phenomenon of activity in the blogosphere we are analyzing is currently an important stream of research that allows you to learn about the next dimensions of the Internet social space.

The novelty of our study is the combination of the two so far separately analysed structural aspects of online communities. In the analyses conducted hitherto, the researchers have either focused on detecting leadership roles characterized by a high degree of influence on others (Johnson et al. (2015)), or focused on examining the level of 140 involvement in the functioning of these communities depending on their roles (Benamar et al. (2017), Rullani and Haefliger (2013)). The purpose of our research is both to discover the influential roles and to show the factors that affect the internal dynamics of online groups. This will allow a more comprehensive and holistic presentation of the mechanisms affecting the structure of the online community.

### 145 2.2. Selected aspects of social media analysis

A popular category of social media are blog portals. They contain *posts* and *comments* written by their users. Posts are usually longer texts concerning specific topics, often having a title and a list of tags written by their authors. Other users may express their opinions regarding posts by writing comments. Comments relate to posts or other comments and are usually much shorter than posts. The comments written by users constitute a social network of interactions

150 between them. Each comment creates a link from its author to its addressee or if such a link exists, increases its weight. Such a network can then be analyzed using algorithms and tools offered by the social network analysis domain (Hanneman and Riddle (2005); Newman and Park (2003); Wasserman and Faust (1994)). Numerous studies on the specifics of online social networks and blogs have been conducted identifying their specific characteristics, such as the existence of well-defined groups. The features of online social networks, such as Small World effect and scale-free property, are analyzed in Heidemann et al. (2012) explaining how they influence a behavior of such  
155 networks.

In Benevenuto et al. (2009), the authors claim that about 80% of interactions concern directly connected friends, which means that usually users of social media have a limited number of constant partners for discussion/interactions, to whom they devote the most attention. Such local interactions play an important role in building social groups. Finding social network groups is fundamental in understanding the properties of the whole network and better understanding human behavior. From a social perspective, a group is a collection of individuals who have been communicating for some time. Individuals forming a group interact or have the ability to interact with any other individual belonging to the group (Homans (1950)). The emerging and existing groups on-line can be defined in the same way (Gritsenko (2016)). What differentiates groups from random sets of individuals is the existence of intra-group structures, based on the exchange of information, which are the effect of adopted roles that translate into the positions held in the group  
160 in terms of impact on its members.

It is difficult to give one accepted formal definition of a group, because of strong relationship discovering groups with the nature of the data and applications (Tennakoon and Nayak (2019)). In most studies (Fortunato (2010), Parthasarathy et al. (2011)), the group is treated as a set of densely connected nodes with very few connections with the nodes not belonging to a given group. There are many division of methods for groups discovery, e.g., finding overlapping or disjoint groups, hierarchical/non-hierarchical, deterministic/non-deterministic. The significant example of the algorithm which extracts overlapping groups is the Clique Percolation Method (CPM) (Palla et al. (2005)), which main relies on the idea that communities consist of several fully and densely connected subgraphs defined as *k-cliques* ( $k$  - number of nodes in the subgraph). A *k-clique community* is defined as the union of all  $k$ -cliques that can be reached from each other through a series of adjacent  $k$ -cliques (two  $k$ -cliques are adjacent if they share  $k-1$  nodes). Other examples of overlapping approaches are DEMON (Coscia et al. (2012), Coscia et al. (2014)) (each node is asked to vote for the communities present in its local view of the network), COPRA (Gregory (2009)) (based on label propagation approach), OSLOM (Lancichinetti et al. (2011)) (based on statistical significance of clusters and hierarchy). Examples of solutions working on non-overlapping groups are algorithms proposed by Girvan and Newman (Girvan and Newman (2002)) or Louvain method (Blondel et al. (2008)) (based on modularity organization and hierarchy).  
170  
175  
180

Recently, there are more and more studies treating the group as a set of nodes with similar interests (Hromic and Hayes (2018), Tennakoon and Nayak (2019)) (semantic based approach, functional communities). A very comprehensive overview of existing methods has appeared and advanced taxonomy was proposed (Azaouzi et al. (2019)). The author recognized the need for distributed processing so the existing approaches divided into two groups: centralized and distributed community detection. In both groups, the algorithms are divided into two additional groups: community detection based only on structural partition and on both structural and semantic partition.  
185

Because of the dynamic nature of social media (Fu (2019)) (users appear and disappear, add their posts and comments, change affiliations to groups, and change their roles in groups), it makes sense to follow the group's life cycle and the changing roles of users belonging to it (Nath and Roy (2019)). In Greene et al. (2010) general strategy of analysis of dynamics of groups was presented. In each time slot the groups are extracted and then groups from neighboring slots are matched (e.g., by calculating Jaccard index between these groups). We proposed (Gliwa et al. (2012b)) Stable Groups Change Identification (SGCI) algorithm which contains some improvements. Important aspect during analyzing the dynamics of groups is the identification of events that occur in the groups lifecycle. The set of events varies in different methods. Palla in Palla et al. (2007) proposed basic events than can appear: growth, contraction, merging, splitting, birth and death. Good survey of the network and groups evolution may be found in Spiliopoulou (2011), Aggarwal and Subbian (2014).  
190  
195

### 2.3. Users influence

In many social media, influential users may affect others, for example by writing posts that are frequently commented on by others and writing comments to other posts. Claims in Stephen et al. (2010) show that people have



200 limited ability to influence others. This is due to the fact that although highly connected people have a big impact on their immediate environment, it is also likely that they can not spend much time communicating with everyone in their neighborhood. As a result, the power of influence of these people is decreasing in the total number of people belonging to their neighborhood (Katona et al. (2011)). This shows that maintaining an influential position in the group requires focusing on building of broad social relations within the group and focusing on a small number of people. This allows to keep the intensity of contacts.

205 Identification of influential users of social media, (e.g., bloggers) becomes more important, as they may have a significant role in formulating political programmes and discussing social issues, and can shape consumer purchasing decisions (Agarwal et al. (2008)). Previous research aiming at identifying influential social media users can be divided into four basic categories.

210 The first one focuses on the identification of influential users of blogging platforms (Agarwal et al. (2008); Akritidis et al. (2009); Cai and Chen (2012)) to discover the factors that make some bloggers being considered influential. The researchers focused on the analysis of blogger's activity in relation to their other characteristics. That allowed us to show the key factors in having a social impact. In the paper (Cai and Chen (2012)) the authors are focusing on features such as the level of recognition of posts and the activity generated by them measured by the number of comments appearing under entries and the number of external links leading to them. To determine influential bloggers, next to the activity generated by entries, the nature of the posts sent, their novelty and eloquence was also taken into account - determined on the basis of the length of entries (Agarwal et al. (2008)). Influential users were also identified by the recognition of their posts determined by the number of comments and external links to their entries in relation to the time in which their entries gained popularity (Akritidis et al. (2009)).

220 The second approach allows the identification of influential network users, taking into account the nature of the content published by the user, highlighting its "novelty" (Song et al. (2007)) or taking into account the two factors of popularity and activity (Eirinaki et al. (2012)). The essence of this research is to link the activity of social network users and the specific characteristics of the spread content, to build models which facilitate such an analysis, e.g., investigating both the similarity of users creating specific groups and the structure of links placed on their accounts (Weng et al. (2010)). A number of research approaches and models used to identify influential bloggers and to explore the factors that enable the influencing networks have been presented in Khan et al. (2017), that analyzed the models for identifying influential network users in the literature, which is not an easy task and requires considering many additional aspects of the analyses.

230 The third research perspective focuses on determining how information spread through social networking sites. Studies in this trend analyze the impact of blogging and social platforms on political and consumer behavior (Dominigos and Richardson (2001); Berry and Keller (2010); Kempe et al. (2003); Gillin (2007)). This also includes research to determine the impact of advertisements placed on social networks (Kenneth and Yang (2011)), as well as studies of the way in which information is distributed in the blogosphere, that can be analyzed by adopting a model of the spread of infectious diseases (Gruhl et al. (2004)). Other works focus on reconstructing psychological factors that could identify influential content by measuring the impact of entries on the intensity of the emotional and cognitive responses of the recipients (McNeill and Briggs (2014)).

235 The fourth trend of research focuses on identifying influential blogs by analyzing which sites and social platforms influence both the outside world and the blogosphere itself (Gill (2004); Anderson (2006)).

#### 2.4. Activity of users

240 Activity is related to the effort a user puts in to get an influential position. In the blogosphere, this effort is to a significant extent related to activity by writing posts and comments. In addition, this effort also includes time spent reading posts and comments, some of which may be later subjects of future comments written, but for the simplicity of the model this factor might be omitted. In various published statistics of social media popularity (e.g. Patel et al. (2011)), activity is often identified with the number of visits of different users on a given portal social networks (Pluciennik (2013)). For our needs a more detailed description of the activity is necessary – considered most significant actions performed on the blog portal. In order to determine it, one can take into account the number of posts and comments written, and perhaps also their characteristics (size, semantic structure, etc.).

245 In Vaca Ruiz et al. (2014) the dynamics of microblog users' activity in relation to the attention that this activity aroused in the community is studied. Activity was described by the numbers of written posts and comments. In Long

250 et al. (2015), the relationship between user activity on the portal and the defined set of metrics describing structural features was examined, while in Benevenuto et al. (2009) various user activity profiles were studied taking into account a set of defined metrics. The research were carried out also to identify clusters of users with similar activity profiles on social networks (Hozhyi and Lamiroy (2017)). Another direction of research, where activity in social media is represented and analysed, is a prediction, the activity was predicted e.g. using a modified logistic regression method (Zhu et al. (2013)). A high activity of bloggers allows them gaining of recognition, which may be also a prerequisite for recognizing an individual as a member of a group (Bion (2003)). With respect to bloggers, the recognition is related to the activity on the portal and the greater it is, the more recognizable the blogger becomes.

### 2.5. Cooperativeness of users

260 Cooperative behavior of given users assumes that they focus on strengthening social relationships through their actions. In this case, bloggers will comment on posts of other bloggers in order to strengthen their position and develop relationships. The opposite way to play a role is *competition*, which leads to focusing on achieving the highest position by the actor. Bloggers post content and comment only on their own posts to achieve their intended individual goal of maintaining an influential position within the group. In this sense, the nature of the activities undertaken by bloggers is due to the structure of communication within the group. Previous analyses show that the communication structure exerts an influence on the way the group operates, and therefore is an important constituent of positions in groups. On-line communication – which may be referred to the way bloggers communicate, recognizing that decentralized content is based on a large number of comments from other users of the portal and builds relationships and satisfaction from such contact, while centralized content serves to maintain a high profile in the blogger group (Gritsenko (2016)).

### 2.6. Roles analysis

270 The concept of *social role* has been the subject of analysis from the beginning of the twentieth century, which underlines the importance of the problem. But there is no single definition, because the meaning of this term is broad and definitions largely dependent on the application. In sociology and psychology, social roles are treated as „cultural objects that are recognized, accepted and used to accomplish pragmatic interaction goals in a community” (Gleave et al. (2009)). In role theory, roles are defined as „those behaviors characteristic of one or more persons in a context” (Biddle (1986)). In turn, in the network analysis, a role is identified as a position that has a distinct pattern of relations to other positions (Wasserman and Faust (1994)). In social media, a definition of a role that seems to be most appropriate treats it as a set of characteristics (relevant metrics) that describe behavior of individuals and their interactions between them within a social context (Junquero-Trabado and Dominguez-Sal (2012)). Now this problem is dealt with by scientists from many disciplines and the methods of machine learning and data mining are included.

280 One can distinguish several approaches for identifying social roles (Zygmunt (2018)). The oldest one is based on equivalence classes (Wasserman and Faust (1994)), where the most appropriate is regular equivalence. Another approach is based on the identification of the core and periphery structure (Borgatti and Everett (2000)) where role is assigned based on membership of a particular area. In approach based on clustering feature vectors, each person is represented by a vector of some of the features that represents its behavior and relationships with the other members of the community and such vector can be clustered, so that people with similar characteristics are placed in one group. Taking into account the character of social media, identifying a role as a set of characteristics that describe behavior of individuals and their interactions between them within social context, seems to be most appropriate (Junquero-Trabado and Dominguez-Sal (2012); Zygmunt (2018)). Recently, in Rossi and Ahmed (2014) an interesting taxonomy of role discovery methods was proposed: graph-based (based on equivalences, computed directly from the graph, without taking into account any features), feature-based (graph is transformed into a feature representation and a role is assigned using node equivalence to a feature representation) and hybrid (combines the advantages of both approaches). The most popular representative of graph-based role discovery method is blockmodel (Doreian et al. (2004)). Forestier et al. (2012) reviewed the approaches to role mining in social network. He categorized the roles on *explicit* and *non-explicit*. *Explicit roles* are defined a-priori and are calculated based on specific criteria and methods. Non-explicit roles are identified by using unsupervised approach (most often clustering) (Labatut et al. (2014)).

290 Due to the fact that today’s societies are organized from many diverse groups, both real and virtual, it is also important to define social role through determining the social position individuals occupy in social networks, in addition to the behavior of individuals, expectations towards them and patterns and principles of fulfilling their roles.



This requires taking into account their affiliation to the groups in which they participate (Abnar et al. (2014)). It also allows us to clarify the dynamics of the individuals social activities and behaviors that are changing under the influence of roles played by individuals, which defines them in social groups in which they participate.

For the first time the roles were considered in the context of groups in Scripps et al. (2007). Four community-based roles have been defined: *ambassadors*, *big fish*, *loners*, *bridges*. The roles are counted based on the approximate number of communities that a node in a network belongs to and the degree of a node relative to its immediate neighbors. However, the dynamics of the groups have not been taken into account.

In Guimera and Amaral (2005) two measures were proposed related to the internal and external connectivity of the node with respect to its community. The first - *within-module degree* - evaluates the relationship between a node and its community in relation to other nodes in the same community. The second - *participation coefficient* - assesses the connectivity of the node with other communities. Both measures were used to characterize the role of the node in the community. By setting thresholds (empirically and permanently) for these measures, seven roles have been defined. This way of setting thresholds has been criticized in Labatut et al. (2014), because in this way the thresholds strongly depend on the nature of the analyzed set. Instead of one measure defining the connection with other communities, three of them have been proposed: *diversity* (evaluates the number of communities to which the node is connected (other than own)), *intensity* (number of external links), *heterogeneity* (the distribution of external links over communities). They use the Louvain Blondel et al. (2008) algorithm to find non-overlapping groups, and k-means algorithm to find the roles by clustering defined measures. Finally, they assigned role names to individual clusters. The approach used still did not take into account the aspect of time and changing groups and roles.

In Henderson et al. (2012), *RoIX* - an unsupervised algorithm based on non-negative matrix factorization was proposed for finding roles. Each node is described as a feature vector, so for  $n$ -node network, there are  $n$  vectors (one for each node) with  $f$  features. For such node-feature matrix, non-negative matrix factorization is used. *RoIX* allow to find similar nodes by comparing their role distributions.

This algorithm was used in the dynamic role changing approach (Rossi et al. (2012), Pei et al. (2018)) for reasons of efficiency, automation and scalability. In Rossi et al. (2012), Rossi et al. (2013), Ahmed et al. (2016) roles were discovered, and then transitions were found between roles while in Pei et al. (2018) a unified model for simultaneous detection of the role and transition of role in the dynamic social networks has been proposed. Transition matrix role is proposed, so roles and transitions are modeled in a unified way. Current and historical views are taken into account: current view - to find roles in the current state of the network, while the historical view teaches the role transitions using the role's past information and the current social network snapshot. In these approaches, roles are analyzed globally, without finding groups in networks.

A small group creates better conditions for cooperative activities, as opposed to a large group (Nosenzo et al. (2015)). There is a tendency to identify the individual with the role and a complementary tendency of the individual to assume the role for the duration of the interaction. This mechanism is referred to as the tendency to allocate roles to individuals (Turner (1968)). Changing roles is possible, which is the result of the interim nature of interactions, where a social process of revising concepts that individuals have regarding the role of their partners is ongoing, as a result of which the role is reaffirmed or revised (Turner (1978)). The results of the analyses show that individuals can lose or gain influential roles. This occurs during interactions, therefore time is very important here. What motivates the identification of the role of a given individual with the role itself by members of a given social group is the flexible relation of the actors with individual roles, which in practice means limited possibilities to change the role, reject one and adopt another (Turner (1978)). Role formation processes also take place in online groups (Welsler et al. (2007)). In this context, the analysis of role dynamics in online groups and factors, that influence it, gains an importance.

In Zygmunt et al. (2018) authors analyzed relations between roles played by the user at the level of entire portal and locally on the given groups and noticed that users with influential roles in the groups frequently play significant roles on the level of the whole portal. In Koźlak et al. (2018) the dynamics of roles in the context of group evolution and taking into consideration different types of events describing reorganisation of roles and how they influence the changes of local roles is analyzed.

The article Abnar et al. (2014) proposes a method of finding roles based on structural properties of social network. It is assumed that the network is dynamic, and composed of disjoint groups changing over time using events defined in Takaffoli et al. (2011). Four roles were defined: *leader*, *outermost*, *mediator* and *outsider*, of which the most important group role is *leader*, and *mediator* is important in the context of communication between groups. Distribution of

350 *closeness* measure is used for identification the most important roles. A small, benchmark Enron email network<sup>1</sup> was used for verification. Users start to be tracked only when they get the *leader* role. In the article, the main focus is on finding *mediators* - the algorithm is described in detail. This article is the closest to our research.

### 3. Conception of the approach

355 The purpose of the presented works is to define a strategy of activities that may indicate how the user should behave on the social networking site, and indirect goals to set upon themselves, to gain or maintain the high influence. The developed model has a set of features that are useful in carrying out this task. It has a well-defined passage of time and shows a representation of the network in subsequent states (snapshots), defining then the structure of groups and roles performed by users. This enables an observation of the network evolution between successive states and examination of roles performed by individual users in groups at that time, identifying whether particular roles have  
360 been maintained or changed.

The general elements of the conception are presented in Fig. 1. The approach used in this paper consists of the following steps:

1. **Division of dataset into time periods** — to analyze dynamics, whole range of time is divided into smaller periods (called here time periods). With each distinguished period, the static network is being built based on  
365 the interactions taking place in the analyzed period. In Fig. 1 (step 1) selected subsequent time periods and network snapshots created for them are presented.
2. **Group discovery in each time period** — having static network in each time period, we can use algorithms of group discovery to find groups in each time period. In this paper, we use the Clique Percolation Method (CPM)  
370 <sup>2</sup> (Palla et al. (2008, 2005)), in which overlapping groups are formed by combining  $k$ -cliques having almost all common vertices. It is also possible to use any algorithm finding overlapping groups. In Fig. 1 (step 2) the identified groups and relations between users within groups are marked with thick lines.
3. **Group dynamics discovery** — in this step, the goal is to find continuations (called transitions) of groups (i.e., we want to match groups from the neighboring time periods). For this case, the SGCI method (*Stable Group Changes Identification*) (Gliwa et al. (2012b); Zygmunt et al. (2011, 2012)) finding stable groups was applied.  
375 Stable groups represent groups which exist in a given time period and have successors in required number of following time periods. Group successor of a given group is a group in the next time period which fulfills assumed set of conditions concerning the percent of users which belong to the both groups. In Fig. 1 (step 3) the connections between successor groups are marked with red dashed lines.
4. **Local role discovery** — user can be a member of multiple groups and, in each group, they can play a different  
380 role. Therefore, the analysis of local roles must be connected with a specific group and the interactions inside such a group must be taken into consideration. Details of the method of role discovery will be described in Section 4. In Fig. 1 (step 4) the roles from the set  $R_1 - R_5$  assigned to given users are shown.
5. **Analysis of local roles** — this step covers the analyses regarding stability of local roles (i.e., whether users  
385 keep their local roles in groups that are their continuations) and changes of such roles. In our example (Fig. 1 (step 4)) the stable roles (users maintaining their roles from previous time periods) have their names written in blue color, and the not stable roles (users which acquired roles different to roles they had in the previous time period) in the red one.

### 4. Proposition of approach to role definition

390 User behavior may be characterized from the point of view of three different important dimensions such as influence, activity and a level of cooperation (presented in Sections 2.3, 2.4, 2.5). In our research, the influence is especially important, because it decides whether a given user affects other users in his neighborhood and because of it may be considered as important. The remaining two dimensions are used to better differentiate the behavior of users.

---

<sup>1</sup><https://www.cs.cmu.edu/~enron/>

<sup>2</sup>[www.cfindex.org](http://www.cfindex.org)

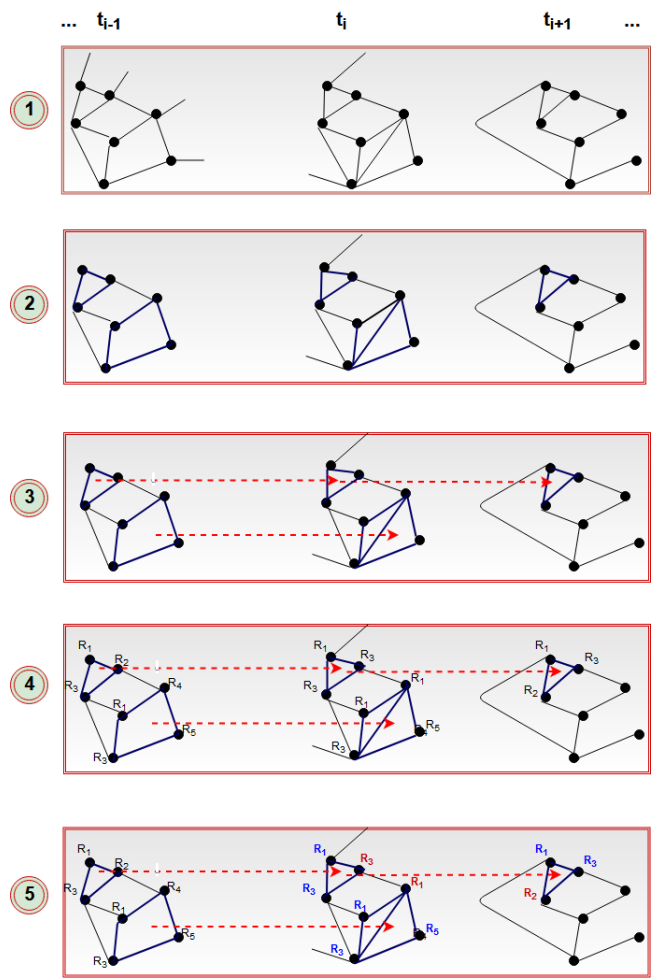


Figure 1: Steps (1–5) in analysis of local roles

Basing the analysis on elements related to the degrees of activity, the cooperation or competitiveness, and influence of the bloggers will allow us to create the typology of roles played by the actors in groups operating in the Internet. By examining changes in groups over time, the dynamics of changes related to the sustainability of specific roles and the degree of transition between actors in groups will be captured.

Defined roles have the task of differentiating vertices in such a way that vertices with similar characteristics and behaviors get to the same role. In this paper we highlighted the following key properties of users in the social media (Gliwa et al. (2016)) based on the concept of role dimensions:

- *Influence* – describes the user’s influence on the community, measures the number of other users’ responses to his/her activity. A high level of influence characterizes the most significant and best-known users of a given portal. In the case of analyzed blog portals, we can also distinguish between the post influence and comment influence. High post influence is particularly important because it means that the user-generated content has a significant perception in the community, attracting many users who want to comment on it. Stable ownerships by the users of high post influence determine that they are ones of the key users of a given portal and knowing them may even significantly go beyond the boundaries of the portal. Big comment influence is less important since the content of comments is generally shorter and less recognizable as they are embedded in the topic context of the post they comment on and the comments made earlier.
- *Activity* – describes the number of actions performed by the user, the number of performed interactions with other users, and is associated with its active engagement in the creation of the content without considering its quality and the scope of impact on other users. It is a particularly important characteristic of a large population of average users that build social portals. In the case of analyzed blog portals, we can distinguish between the user’s activity related to writing posts and/or writing comments. Post activity is a prerequisite to achieve a high degree of recognition in the groups. Comment activity serves mainly to sustain social relationships and the recognition of individual users.
- *Cooperativeness* – describes additional characteristics of users and is particularly important for interpreting the behavior of influential people. The user can either limit their activity to respond to the action taking place in the context of his/her posts, or be willing to respond to the posts written by other users and the comments made in the context of these posts. This activity builds social relations on the portal and strengthens the importance of other users by increasing their recognition and possibly stimulates discussion that may further enhance this recognition. Such users can be classified as social, in contrast to those users who only build their position (selfish).

In the process of assigning roles to users, several fundamental parameters are used:

- *isPostInfluential* – determines whether a given user is considered an influential author of posts, it is related to writing a sufficiently high number of posts considered influential, a post is considered influential if it has been commented by an sufficiently high number of comments, which authors have high post influence.
- *isCommentInfluential* – determines whether a given user is considered an influential author of comments, it is related to writing a sufficiently high number of comments considered influential, a comment is considered influential if it has been commented by an sufficiently high number of comments, which authors have high comment influence.
- *selfishRatio* – determines which part of all comments of the user is written in the context of his posts (and thus including comments on his posts and subsequent comments on comments to them), when assigning roles, it considered whether the value *selfishRatio* is greater or smaller than the *selfishRatioTh* threshold.
- *writtenPosts* – number of posts written by the user, when assigning roles, it is taken into account whether the user has written more or fewer posts than the average number of posts written per user (*postNoMean*).
- *writtenComments* – number of comments written by the user, when assigning roles, it is taken into account whether the user has written more or fewer comments than the average number of comments written per user (*comNoMean*). We assume that the number of comments written by users having influential social roles is not too low, so it should be higher than a defined threshold *writtenCommentsTh*.

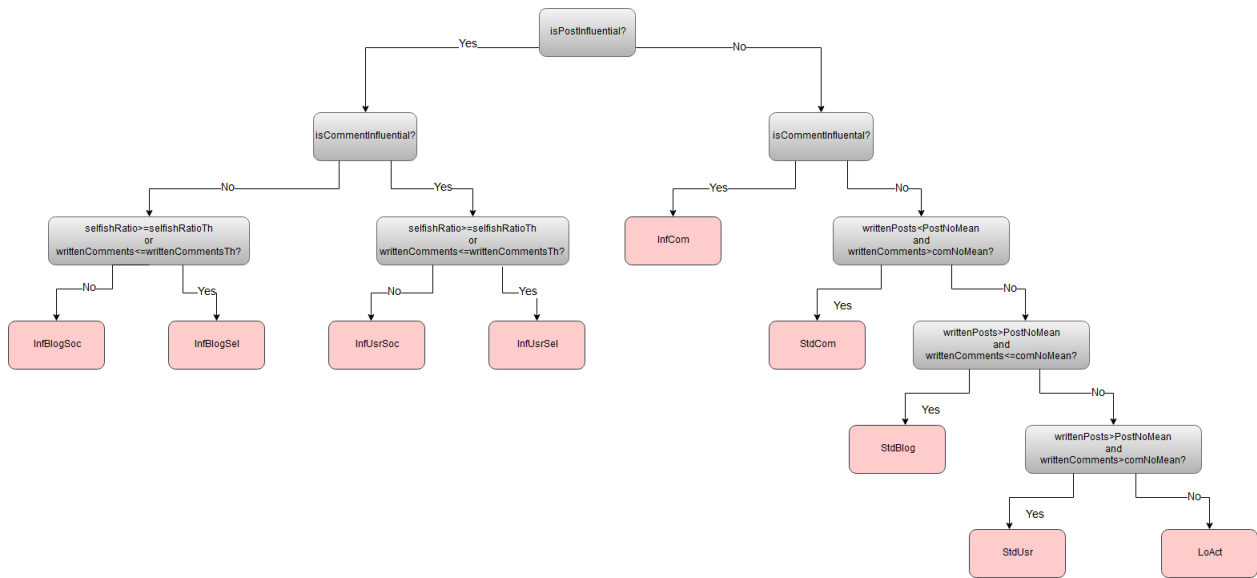


Figure 2: Relations between roles

440 The definition of roles are based on functions that individual perform within a group (Hare (1994)). Given the above characteristics, we distinguish between the following roles (the relations between roles with connections to parameters described above are presented in Fig. 2):

- 445 • *Influential User Social (InfUsrSoc)* – highly influential users (related to writing posts and comments), social in nature, who write a significant number of comments also to the posts of other users,
- *Influential User Selfish (InfUsrSel)* – highly influential users (related to writing posts and comments), selfish in nature, who write comments primarily in the context of their own posts,
- *Influential Blogger Social (InfBlogSoc)* – highly influential users (associated only with writing posts), social in nature, who write a significant parts of the comments also to other users’ posts,
- 450 • *Influential Blogger Selfish (InfBlogSel)* – highly influential users (related only to posts), selfish, who write comments primarily in the context of their posts,
- *Influential Commentator (InfCom)* – highly influential users associated with writing comments,
- *Standard Commentator (StdCom)* – users who write more than the average number of comments and less than the average number of user posts,
- 455 • *Standard User (StdUsr)* – users who write posts and comments in quantities greater than the average per user, but the responses to them are not high enough to have them included in one of the influential roles.
- *Standard Blogger (StdBlog)* – users who write less than the average number of comments per user, but more than the average number of posts,
- *Low Active (LoAct)* – users who write both posts and comments in amounts less than the calculated mean values for the considered active bloggers on the portal.

460 Roles describe the meaning and behavior of the user in the community. Different types of communities can be explored on the blog portal: either globally all users, or locally within the separate communities (groups) strongly

linked together. Our previous works mainly relate to global roles (Gliwa et al. (2013), Gliwa et al. (2016)), in this paper we focus on the local roles, in particular groups.

465 If we deal with a significant group (taking into account its size or members), the user performing an important role in the group may gain importance in a wider circle, without necessarily being qualified for a significant role at the whole portal level. So the group has an impact on the chances of the user to maintain an important role for a long time. Observation and analysis of such dependencies is one of the important goals of our work.

## 5. Model of social system for analyzing users behavior

470 The aim of the defined model is to analyze the behavior of given users in groups of a social portal and to help distinguishing classes describing different behaviors as well as to define notions, which are useful in characterising influential users, that maintain their important position, loose it, or users who are promoted to important roles. These notions are then used to describe and interpret experiments presented in Section 6.

475 In Gliwa et al. (2013), a first model of a social system has been introduced. It was characterized by three elements: *global roles* played by given entities in the whole network, *stable groups* and *local roles* assigned to entities in each stable group. Each user at any given point in time may be described by the role played globally and by roles played in each group it belongs to. However, so far, the main attention was paid to the analysis of global roles. We have modified our approach and now propose another version of the model to better understand and describe the whole system, its components and dependencies between them. The model of the social system  $S$ , is represented by the following  $n$ -tuple:

$$S = (U, L, SG, R, \rho, \tau) \quad (1)$$

480 where

$U$  – set of users  $u$  ( $u \in U$ ),

$L$  – set of links/relations  $l$  ( $l \in L$ ),

$SG$  – set of stable groups  $sg$  ( $sg \in SG$ ),

$R$  – set of roles  $r$  ( $r \in R$ ),

485  $\rho$  – function assigning roles to users in given stable groups (see Equation (8)),

$\tau$  – function of role transition describing changes of roles of a user belonging to the given stable groups between two subsequent time periods where the group exists (see Equation (11)). The function returns a pair  $(r_a, r_b)$ , where  $r_a$  is a role of the user in the considered time period, and  $r_b$  is the role of this user in the next time period in the same stable group.

490 We can represent the social system in time  $t$  as  $S(t)$ . To describe it, for each time period  $t$ , we introduce subsets of the previously defined sets  $U$ ,  $L$  and  $SG$  named  $U_t$ ,  $L_t$  and  $SG_t$  expressing the sets of users, links and stable groups existing in this time period respectively.

Stable group  $k$  which exists between  $t_i$  and  $t_j$  time slots (represented as  $sg_{t_i, t_j}^k$ ) is a sequence of the stable group compositions  $sgc_t^k$  for  $t_i \leq t \leq t_j$ :

$$sg^k = sg_{t_i, t_j}^k = ((sgc_{t_i}^k, t_i), \dots, (sgc_{t_j}^k, t_j)) \quad (2)$$

495 Stable group composition of group  $k$  at time  $t$  ( $sgc_t^k$ ) is a set of users belonging to the stable group  $k$  at this time. For these users one can assign also roles played by the users in this group at this time. For a stable group  $sg^k$  existing at time  $t$  one can test whether a given user  $u$  belongs to it at this time, or not. In a given time period  $t$  a user  $u$  may belong to a group  $sg^k$  or may be outside this group:

$$u \in U \wedge (\forall sg^k \in SG_t : u \in sg^k \vee u \notin sg^k) \quad (3)$$



A user can play one role (from the  $R$  set) in a given stable group and in a given time period  $t$  ( $t_i \leq t \leq t_j$ ).

$$\forall u, sg_{t_i, t_j}^k, t_i \leq t \leq t_j \ u \in sg^k \Rightarrow \exists! r \in R : r = \rho(u, sg_{t_i, t_j}^k, t) \quad (4)$$

500 The set  $R$  contain all user roles described in Section 4.

$$R := \{InfUstrSoc, InfUstrSel, InfBlogSoc, InfBlogSel, InfCom, StdCom, StdBlog, LoAct\} \quad (5)$$

Among roles in groups we distinguish between two sets of important roles:  $I_1$  (which contains roles of users with influential posts) and  $I_2$  (which contains  $I_1$  roles extended by  $InfCom$  role).

$$I_1 := \{InfUstrSoc, InfUstrSel, InfBlogSoc, InfBlogSel\} \quad (6)$$

$$I_2 := I_1 \cup InfCom \\ = \{InfUstrSoc, InfUstrSel, InfBlogSoc, InfBlogSel, InfCom\} \quad (7)$$

Function  $\rho$  assigns a role  $r$  to user  $u \in U_t$  in a stable group  $sg_{t_i, t_j}^k \in SG_t$  for a time period  $t$  (where  $t_i \leq t \leq t_j$ ).

$$\rho(u, sg_{t_i, t_j}^k, t) = \begin{cases} r : r \in R & \text{if } u \in sg_{t_i, t_j}^k \\ \emptyset & \text{if } u \notin sg_{t_i, t_j}^k \end{cases} \quad (8)$$

For each stable group composition  $sg_{t_i, t_j}^k$  we define *the densities of the important roles* functions  $\varphi_1$  and  $\varphi_2$ :

$$\varphi_1(sg_{t_i, t_j}^k) = \frac{\#(I_1(sg_{t_i, t_j}^k))}{\#(sg_{t_i, t_j}^k)} \quad (9)$$

$$\varphi_2(sg_{t_i, t_j}^k) = \frac{\#(I_2(sg_{t_i, t_j}^k))}{\#(sg_{t_i, t_j}^k)} \quad (10)$$

505 where:

- $\#(I_1(sg_{t_i, t_j}^k))$  – number of users having a role belonging to  $I_1$  set in group composition  $sg_{t_i, t_j}^k$ ,
- $\#(I_2(sg_{t_i, t_j}^k))$  – number of users having a role belonging to  $I_2$  set in group composition  $sg_{t_i, t_j}^k$ ,
- $\#(sg_{t_i, t_j}^k)$  – number of all users in group composition  $sg_{t_i, t_j}^k$ .

510 Role transition ( $\tau$ ) describes the change of the role of the user  $u$  between the two subsequent time periods  $t_j$  and  $t_{j+1}$ , where this group exists and  $u$  belongs to it.

The following assumptions are made:

- $t_i$  and  $t_m$  are first and final time periods of existence of stable group  $sg_{t_i, t_m}^k$ , where  $t_{j+1} \leq t_m$ ,
- $j$  is a time period when the group  $sg_{t_i, t_m}^k$  exists and is not a last time period of its existence,

Then we define the role transition function ( $\tau$ ) as:

$$\begin{aligned} (u \in sg_{t_j}^k \wedge u \in sg_{t_{j+1}}^k) &\Rightarrow (\tau(u, sg_{t_j, t_m}^k, t_j)) \\ &= (\rho(u, sg_{t_j, t_m}^k, t_j), \rho(u, sg_{t_j, t_m}^k, t_{j+1})) \end{aligned} \quad (11)$$

515 One can distinguish between different classes of transitions which have different features and different levels of significance.

Conditions for important transitions between  $t_j$  and  $t_{j+1}$  time periods are as follows:

- the same role is preserved:

$$\tau(u, sg_{t_j, t_m}^k, t_j) = (r_a, r_b) : r_a = r_b, r_a, r_b \in R \quad (12)$$

- important role (from  $I_1$  set) is preserved:

$$\begin{aligned} \tau(u, sg_{t_j, t_m}^k, t_j) = (r_a, r_b) : r_a \in I_1 \wedge r_b \in I_1, \\ r_a, r_b \in R \end{aligned} \quad (13)$$

- important role (from  $I_1$  set) moves to not important one:

$$\begin{aligned} \tau(u, sg_{t_j, t_m}^k, t_j) = (r_a, r_b) : r_a \in I_1 \wedge r_b \in R \setminus I_1, \\ r_a, r_b \in R \end{aligned} \quad (14)$$

- 520
- not important role moves to important one:

$$\begin{aligned} \tau(u, sg_{t_j, t_m}^k, t_j) = (r_a, r_b) : r_a \in R \setminus I_1 \wedge r_b \in I_1, \\ r_a, r_b \in R \end{aligned} \quad (15)$$

Similar conditions may be introduced taking into consideration the belonging of the user roles to the  $I_2$  set.

525 The proposed model is generic and may be extended to describe different kinds of system evolution (for example detailed conditions for stable group identification, different roles sets, different kinds of patterns identifying role transitions, group subsets or group evolution). The elements of the model are used to describe experiments presented in Section 6.

## 6. Experiments concerning role distribution and transitions

The experiments described in this section take advantage of the model presented in Section 5. The main goal is to verify which user behaviors give them the higher chance of gaining and maintaining a significant position. For this purpose we examine the character of two popular blog portals.

### 530 6.1. Aims of analyses

The objective of the work is to explore transitions of user roles (different values of  $\tau$  function defined in Equation (11)). Particularly, we are interested in frequency and conditions of appearance of these transitions taking into consideration the important roles related to a post (roles in the  $I_1$  set, defined in Equation 6) and comment (some roles from  $I_1$  and role which appear additionally in  $I_2$ , which is defined in Equation 7).

535 In relation to this the following is examined:

- which portion of users playing given roles in groups remain in groups as a consequence of transition (assuming that a group has a significant influence on building user position) (Section 6.6.1),
- which portion of users having given roles maintain their roles after the transition (Section 6.6.2),
- which portion of users with important roles maintain the important roles in the next time periods (Section 6.6.2),
- 540 • which portion of users with unimportant roles promote to important roles ( $I_1$  or  $I_2$ ) in the next time periods (Section 6.6.3),
- which portion of users lose important roles (Section 6.6.3),
- which other transitions between given pairs of roles appear frequently (Section 6.6.4).

In order to properly verify the obtained results, we present characteristics describing the features of portals and the specificity of identified user groups and roles:

- the numbers of groups with given sizes and numbers of roles (Section 6.3),
- the density of important roles in groups ( $\varphi_1$  and  $\varphi_2$  functions defined in Equations (9) and (10) ) and depending on their sizes (Section 6.4),
- the numbers of roles in groups with considered sizes, to verify whether group sizes have influence on the role distribution and in which way (Section 6.5).

## 6.2. Datasets description

The data from the two national blogospheres: Polish blogosphere *Salon24*<sup>3</sup> and American blog portal *The Huffington Post*<sup>4</sup> are analysed. Political blogs constitute the largest part of *Salon24* portal while *The Huffington Post* contains news and blogs from various subjects, where political topics constitute a significant part of all posts, but this topic does not outnumber other ones as in the case of *Salon24*.

Both portals have different characteristics (Zygmunt and Gliwa (2015)), they differ in how the users use them. There is a strong distinction to be made between posts writers and commentators in *The Huffington Post*, where they constitute in reality two separate groups, while there is no such distinction in *Salon24*. The differences in the characteristics should also be reflected in the groups found and roles performed by users in the groups.

The data from *Salon24* is from a time range of 1.01.2008 – 6.07.2013 and contains 380,700 posts, 5,703,140 comments and 31,750 authors. The second dataset, *The Huffington Post* contains data from a period of 1.01.2010 – 14.11.2013 and embraces 414,225 posts, 17,796,819 comments and 680,341 authors. Both datasets have similar numbers of posts, but in *The Huffington Post*, there are three times more comments and twenty-five times more authors. In both cases, the whole period of time was divided into overlapping time periods (called here *time slots*), each lasting 7 days, and the neighboring slots overlap each other by 4 days. In this way, the dataset from *Salon24* contains 504 slots, and from *The Huffington Post* – 345. In each time slot, the static networks were built in the same way for both data sets (*comment model* described in Gliwa et al. (2012a)): users are nodes and relations between them were built based on comments written.

## 6.3. Characteristic of portals in the context of discovered groups and number of roles

The number of stable groups  $SG$  (element of model described in Section 5), identified in all considered time slots with their sizes in both datasets, is presented in Fig. 3 (Zygmunt and Gliwa (2015)). One can observe that *The Huffington Post* contains more groups overall and this occurs for all group sizes except for groups larger than 100 members. *Salon24* has clearly more largest groups with a size larger than 200 members.

In each group, local roles (belonging to the  $R$  set), described in Section 4, were discovered based on conception presented in Sections 3 and 4. Table 1 shows the comparison of distribution of all discovered roles in groups in

<sup>3</sup>www.salon24.pl

<sup>4</sup>www.huffingtonpost.com

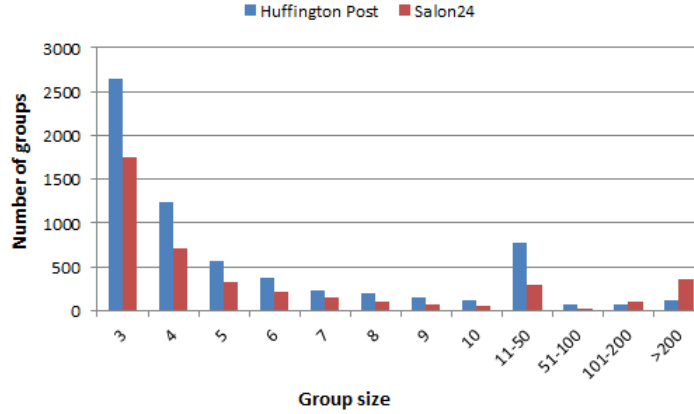


Figure 3: Number of stable groups  $sg^k$  at given size

Table 1: Comparing percentage of roles (in %) in groups for *Salon24* and *The Huffington Post*

Role	<i>Salon24</i> [%]	<i>The Huffington Post</i> [%]
<i>StdBlog</i>	14	8.8
<i>InfBlogSel</i>	3.67	4.9
<i>InfBlogSoc</i>	1.5	0.0002
<i>StdCom</i>	13.12	31.24
<i>InfCom</i>	6.82	9.6
<i>LoAct</i>	46.93	45.47
<i>InfUsrSel</i>	1.57	0
<i>InfUsrSoc</i>	1.25	0
<i>StdUsr</i>	11	0

both datasets: *Salon24* and *The Huffington Post*. In the analyzed period, in *Salon24*, 198,001 roles in groups were found. Four most important, influential roles ( $I_1$ ) constitute 7.99% of all roles. *InfCom*, a relatively less important role because it is only related to receiving comments, constitutes 6.82%. So, the most important and fairly important roles ( $I_2$ ) account for about 15% of all roles, almost as much as *StdBlog* (14%), one of the roles with ordinary importance.

580 The structure of roles in *The Huffington Post* is quite different: more roles in groups were found (245,286), but the most influential ones, that are *InfUsrSoc*, and *InfUsrSel* (belonging to  $I_1$ ) were not found at all (Table 1). Considering the much larger number of *The Huffington Post* users and the similar number of roles in the groups identified for both sites, significantly fewer *The Huffington Post* users have been qualified for discovered groups. This is connected to different structures of the two portals. If users write posts on *Huffington Post*, they very rarely comment on them as  
585 opposed to *Salon24*, where there is no such distinction between users. Also in the *The Huffington Post* portal many users write their comments only very irregularly, addressing them only for particularly interesting posts.

This is also reflected in the diversity of the number of important bloggers: in *The Huffington Post* they are much less interested in commenting on posts not created by themselves than in *Salon24* (which means more *InfBlogSel* and far fewer *InfBlogSoc*). In *The Huffington Post* there is a small group of post writers and much a larger one of  
590 commentators, which is reflected in a significantly greater number of commentators (*StdCom* and *InfCom*). Both services have a comparable number of the low active users, who activate incidentally from time to time.

#### 6.4. Density of important roles in groups with different sizes

Because there is a difference in the number and size of the groups, as well as in the number and the structure of roles in both portals (described in Section 6.3), we try to verify whether the density of important roles ( $\varphi_1$  and  $\varphi_2$ )  
595 depends on the size of the group.

Analyzing the density of important roles ( $\varphi_1$  and  $\varphi_2$ ) in both *Salon24* and *The Huffington Post* datasets (Fig. 4 and Fig. 5), it can be concluded that the shapes of functions for both portals are similar. The differences are that in case of

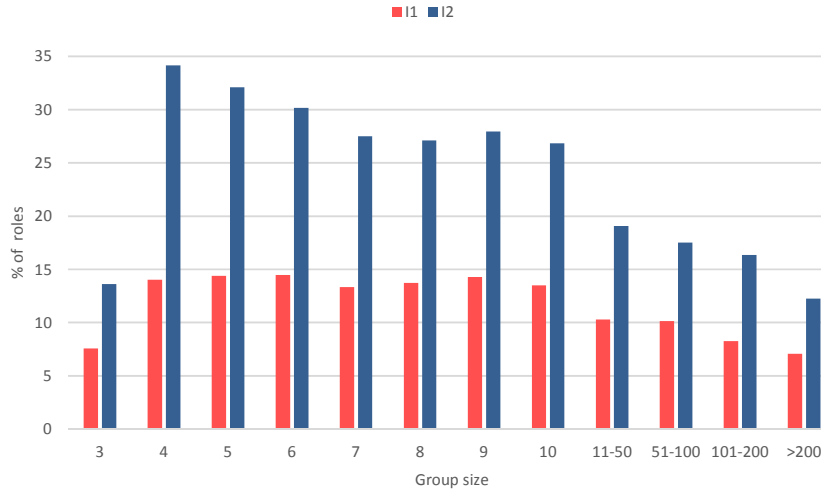


Figure 4: Density of important roles  $\varphi_1$  and  $\varphi_2$  for different groups sizes in *Salon24*

*The Huffington Post*, the value of  $\varphi_2$  is higher than for *Salon24* and the value of  $\varphi_1$  is lower than for *Salon24*. This is related to differences between portal characteristics, significantly more frequent presence of *InfCom* role in groups on *The Huffington Post* portal, especially considering it with the relation to the frequencies of significant roles belonging to  $I_1$  set on given portals.

In both datasets the density of important roles  $\varphi_1$  contains significantly lower percentage of users than  $\varphi_2$ , differences are especially large for groups with sizes between 4 and 10. The highest value of  $\varphi_2$  is for groups with size 4, for higher sizes of groups this value decreases.

Generally, density of important roles in groups with given sizes is relatively high. In more details,  $\varphi_1$  differs between slightly below 15% for *Salon24* and slightly below 10% for *The Huffington Post* for groups with sizes from 4 to 10. An even higher value is obtained for  $\varphi_2$  (which means that this increase is caused by users with influential commentator role *InfCom*), which reaches between 25% and 35% for both *Salon24* and *The Huffington Post*, and the groups with sizes from 4 to 10, the same period as the one considered for  $\varphi_1$  function.

The significant frequency of users with influential roles in groups is associated with the fact that to the groups belong only users having relevantly strong and stable connections with other users at least within the given group.

### 6.5. Dependence between sizes of groups and distributions of roles

When analyzing the roles from  $I_1$  in both datasets in more details (Figs 6 and 7), we can notice that *InfUsrSel* and *InfUsrSoc* roles have low frequency in majority of groups sizes in *Salon24* and do not appear at all in *The Huffington Post*. *InfBlogSel* reaches the highest frequency for rather small groups (with sizes 5–8) and its value strongly decreases for higher groups. *InfBlogSoc* is very low in *Salon24* and almost absent in *The Huffington Post*. The results of the analyses show that roles oriented toward building ones own position (role type *selfish*) are the most frequent in small groups. This confirms the results of research showing that it is easier to achieve an important selfish role in small groups (Hirvonen (2016)). The existence of cooperation oriented influential users favours the emergence of larger groups, hence their weaker representation in small groups.

We can see that *InfBlogSoc* is rare in all studied groups, this indicates that groups which form on blogs are dominated by competing rather than cooperating roles (Tab 1 shows that selfish roles are more numerous than their social counterparts in both *InfUsrSel/InfUserSoc* and *InfBlogSel/InfBlogSoc* cases). This is because acquiring recognition while expressing social and cooperative behavior requires actors to have a high level of involvement in the role. They must post and comment often if they want to achieve and maintain influential positions. In addition, as published research has shown, posts must be up-to-date and/or popular (Song et al. (2007); Eirinaki et al. (2012)), which requires analyzing messages that appear in other media, to attract other users, attention and to achieve a high level of recognition, which translates into maintaining an influential role. As a result, bloggers more often focus solely on

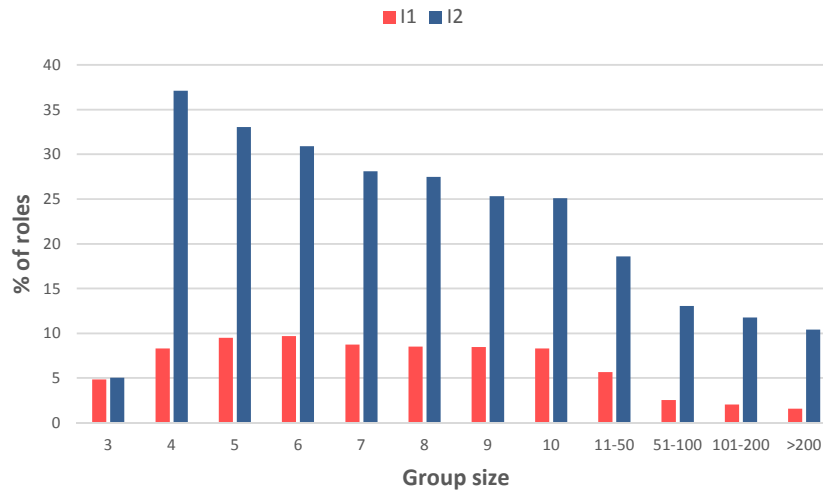


Figure 5: Density of important roles  $\varphi_1$  and  $\varphi_2$  for different groups sizes in *The Huffington Post*

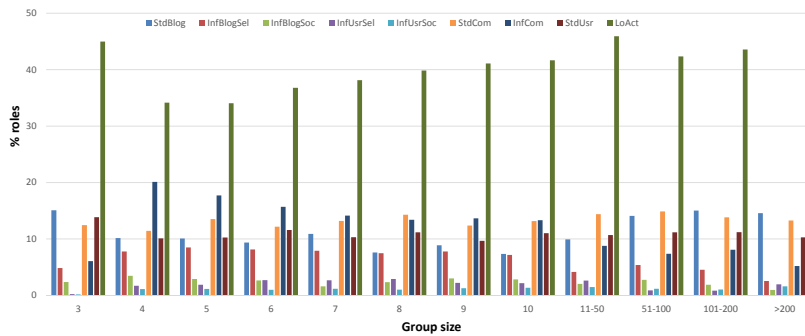


Figure 6: Percentage of roles in different groups sizes in *Salon24*

630 achieving their own goal of gaining an influential position, high in the power structure of the group, without engaging  
 in building and maintaining relationships with other users. Additional important role from  $I_2$ , *InfCom*, often appears  
 in both datasets in small groups, while in *The Huffington Post* it has a significantly higher frequency.

The presented analysis indicates the relationship between the size of the group and the roles played by the users,  
 so the character of the group measured by the number of members contributes to the frequency of influential roles  
 assigned to users. This shows that as the group grows larger, voluntary engagement in the production of goods  
 635 common to the group weakens (in this case: posts and comments). In addition, as the number of members grows,  
 the frequency and quality of communication decreases, and the group's homogeneity weakens, making adopting  
 influential roles difficult. This explains why more actors have influential positions in small and medium groups  
 (Hirvonen (2016)). The only roles which frequencies are independent from group sizes are *StdCom* and *StdUsr* in  
*Salon24*, while in *The Huffington Post* it is true for *StdCom* in the groups with sizes larger than 3.

640 The behavior of the users on both social media portals is different. Users of *Salon24* often write both posts and  
 comments while on *The Huffington Post* users are well separated into two group – publishing posts and commenting  
 them. In *Salon24* interactions between users have more frequently a mutual (two sided) character. As a result, roles  
 are less diverse in *The Huffington Post* than in *Salon24*.



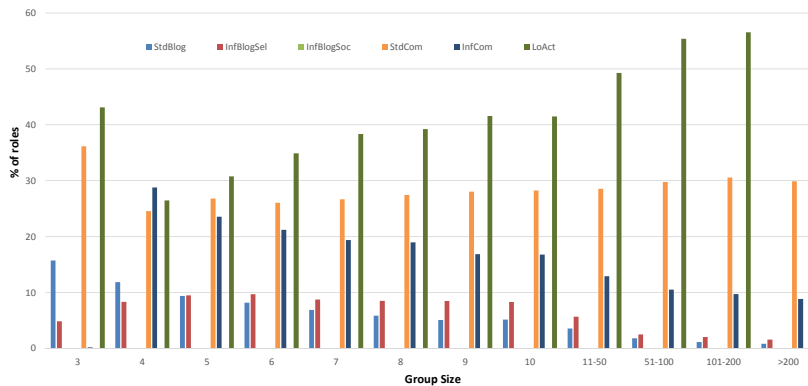


Figure 7: Percentage of roles in different group size in *The Huffington Post*

## 6.6. Transition of roles between groups

645 A user performing a specific role in a group at a given time slot can, during transition, remain in a group while  
maintaining the role or changing it to another (with lower or higher significance), it may also no longer remain in this  
group. It could also happen that a user does not belong to any group. This was discussed in the model description  
of  $\tau$  function (Equation (11)) and its different forms (Equations (12)-(15)). In *Salon24* 109,410 roles participate in  
transitions in consecutive slots, while in *The Huffington Post* 48,049 users participate, which is less than half of the  
650 number for *Salon24*.

You can see that these values are clearly lower in the case of *The Huffington Post*, which results from the smaller  
regularity of writing posts by influential users of *The Huffington Post*, than it is in the case of *Salon24* users. If there  
are so many users with assigned roles which remain in the group in the next slot, it is worth checking which roles  
played by users are stable and which are not.

### 6.6.1. Roles remaining in groups during transition

655 The aim of the experiments described in this section is to check which percentages of each kind of role remain in  
group after transitions (Tab. 2). Especially interesting roles are those that belong to  $I_1$  or  $I_2$  sets. In both considered  
datasets the most influential roles in the majority of cases still belong to groups. Generally, in the case of *The*  
*Huffington Post* portal, significantly smaller percentages of users remain in groups.

660 In *Salon24* (Tab. 2), the roles which most frequently remain in groups during the next time slot are the most  
important ones: *InfUsrSoc* and *InfUsrSel* (95% and 90% respectively) as well as *InfBlogSoc* (95%). Users having  
*InfBlogSel* role much less often remain in the group (in about 77% of cases), it is the consequence of frequently  
weaker connection of such users with other users belonging to the group and significantly less systematic writing of  
posts by them. In the case of *The Huffington Post*, the highest percentage of roles which remain in groups are for  
665 *InfBlogSel* (almost 64%), and *InfCom* (55%) which are the most frequent important roles in this dataset.

### 6.6.2. Maintaining roles during transitions

Subsequent analysis concerns a verification of which portion of users, which remain in their groups in the next  
time slots (described in Section 6.6.1), preserves the roles they had in the previous time slots.

670 In Tab. 3 one can see the percentage of transitions to the group with preserved roles (column *Keep role*) in the  
next time slot or those that have the role that belongs to important influential roles in sets  $I_1$  or  $I_2$  (it is described in  
the next Section 6.6.3) on the next time slot. In table cells, the first number concerns *Salon24* and the second one *The*  
*Huffington Post*.

675 When analyzing what percentage of users have maintained their roles in *Salon24* (Tab. 3 column *Keep role*),  
a significant part of roles from  $I_1$  maintain their roles, but when comparing percentages of users with given roles  
maintaining their roles in both datasets one can see that the part of preserved given roles is higher in *The Huffington*  
*Post*. These higher percentages may be partly explained by the fact, that some roles do not appear in the *The Huffington*

Table 2: Percent of roles remaining in a group in consecutive slots

Role	Salon24	The Huffington Post
<i>StdBlog</i>	60.17	55.53
<i>InfBlogSel</i>	77.20	63.56
<i>InfBlogSoc</i>	95.01	—
<i>InfUsrSel</i>	90.59	—
<i>InfUsrSoc</i>	95.01	—
<i>StdCom</i>	85.63	48.70
<i>InfCom</i>	81.60	55.30
<i>StdUsr</i>	87.33	—
<i>LoAct</i>	55.44	39.48

*Post*, so the choices of changing roles to another ones are limited. In the case of *Salon24* the highest stable roles from the  $I_1$  set are: *InfUsrSel*, *InfBlogSel* and *InfUsrSoc*, while in *The Huffington Post* they is *InfBlogSel*. Relatively many *InfCom* roles (from  $I_2$  set) maintain their role. Especially, many users (78% and 80% in *Salon24* and *The Huffington Post*) with low important *LoAct* role keep their roles, which means that users with activities below average usually remain on that low level.

Concluding then, one can observe a high stability of important roles belonging to  $I_1$  or  $I_2$  in both datasets. For users with less important roles, if they remain in groups, they do not change their behavior in their groups and more often maintain their roles than change them. In general the stability of important roles is lower than the stability of not important ones.

### 6.6.3. Promotion and demotion of roles

The next research problem is to verify which roles give the highest chance to maintain an important role or to promote to such a role by a given user, which means that a user either keeps an influential (belonging to  $I_1$  or to  $I_2$ ) role or changes the role to one of the influential roles.

An analysis of the probability of obtaining important roles belonging to  $I_1$  can also provide useful information (Tab. 3, column  $I_1$ ). For *Salon24*, a significant part of the important roles remains important (still belonging to  $I_1$ ) – it is true for *InfUsrSel* (71.79%), *InBlogSel* (61.95%), *InfUsrSoc* (61.42%) and *InfBlogSoc* (45.16%). So, among the most influential roles, the *InfUsrSel* is the most stable role and the *InfBlogSoc* is the least stable. In the case of *The Huffington Post*, the only influential role that is frequent and has a significant chance of maintaining its influential position is the role of *InfBlogSel*.

One can see that those roles oriented on building their own influence in the context of their own posts (*selfish roles*) are more stable than the more cooperative roles. Influential roles of *User* type, are more stable than *Bloger* type roles. It is a consequence of smaller dependency on activities of other users, for example, posts with interesting topics published by them and the fact that the communities of users strongly connected with a given user by intensive commenting on their posts create stable connections. The chances of being promoted to influential roles (belonging to  $I_1$  set) are quite significant for *InfCom* on the *Salon24* portal - 8.70 %, there is no similar chance on *The Huffington Post*.

In *Salon24*, among less influential roles, only users with *StdUsr* role (8.96%) and *StdBlog* role (6.84%) have some chances to move up to significant roles. In the case of *The Huffington Post*, the users with the role of *StdBlg* (24.52 %) have even a much greater chance of being promoted to an influential role (24.52%).

We will now consider promotions of non-significant roles to the role of *InfCom* (which is the difference between the sets of  $I_2 - I_1$ ). In *salon24*, quite a few chances (over 7.5%) for such promotion have the roles *StdCom* and *StdUsr*, and so the roles have a significant (above average) number of written comments. Thanks to their quantity, they manage to achieve the necessary level of their influence to be promoted to this role. In *The Huffington Post*, the situation is similar: *StdCom* has a significant chance (about 10%) to be promoted to the role of *InfCom*.

Demotion of users, considering their influence, can be classified into two categories: demotion of influential users (belonging to the  $I_1$ ) and demotion of users having the role of *InfCom*. In the first case, for *Salon24* influential roles, the highest risk of demotion (falling out of  $I_1$ ) is for *InfBlogSoc* role (55%). In *The Huffington Post* the probability of demotion of the only influential role on this portal (*InfBlogSel*) - 41%, is slightly higher than a chance of demotion

Table 3: Percentage of stable transition from given role to the same role (*Keep role*) or to one of influential roles belonging to  $I_1$  or  $I_2$  in groups in next slot *Salon24/The HuffingtonPost*

from role	<i>Keep role</i>	$I_1$	$I_2$
<i>StdBlog</i>	60.96/74.89	6.84/24.52	9.74/24.52
<i>InfBlogSel</i>	46.85/59.05	61.95/59.05	64.26/59.05
<i>InfBlogSoc</i>	24.60/—	45.16/—	51.07/—
<i>InfUsrSel</i>	49.10/—	71.79/—	79.08/—
<i>InfUsrSoc</i>	44.27/—	61.42/—	76.02/—
<i>StdCom</i>	59.70/72.31	1.94/0	9.70/10.10
<i>InfCom</i>	39.16/50.60	8.71/0	47.87/50.60
<i>StdUsr</i>	57.82/—	8.96/—	16.60/—
<i>LoAct</i>	78.59/81.22	1.66/0.12	4.90/5.73

715 for the same role in *Salon24* – 38%. In the second case, the role of *InfCom* has a similar chance of demotion to the usual role on both portals (about 50%).

#### 6.6.4. Frequent transitions between pairs of roles

In Section 6.6.3, only promotions to important roles ( $I_1$  or  $I_2$ ) were analyzed. Now we analyse not only these promotions, but the frequent transitions between all considered roles. In Tabs 4 and 5, the transitions between given pairs of roles for both datasets are considered. In the rows, the percentages of transitions to other roles for given source roles are described. The names of source roles are in the column labeled *from role*, in subsequent columns, the results for given destinations roles are presented (labels *to role*). The last column (labeled *in groups*) describes the portion of roles that remain in groups after transition (described in detail in Section 6.6.1), remaining owners of given roles do not belong to any groups in the subsequent time slot.

725 Frequent transitions between roles can be considered from the perspectives of three highlighted dimensions describing the roles: activity, influence, and cooperativeness/competition. To characterize the nature of a given frequent transition, we decided to introduce the following designations:  $i$  – influence, and  $a$  – activity,  $c$  - cooperativeness, the indexes  $p$  and  $c$  used for  $i$  and  $a$  mean the post and comment influences and activities respectively, and  $(+/-)$  characters mean acquisition or loss of a given feature. These symbols are used in the text below and in Fig. 8, where the characteristics of frequent transitions on the analyzed portals are given.

730 Frequent transitions (marked in bold in Tab. 4) for *Salon24* portal for each input roles are as follows:

- from *StdBlog*:
  - to *StdUsr* – increase of the total number of comments to above the average ( $a_{c+}$ ),
  - to *LoAct* – decrease of the total number of posts to below the average ( $a_{p-}$ ),
- 735 • from *InfBlogSel*:
  - to *StdBlog* – lose post influence ( $i_{p-}$ ),
  - to *StdUsr* – lose post influence, keep writing many posts and comments, not guaranteeing preserving significant post influence, because of not interesting subjects of posts, group reorganization or disappearing the users, commenting on their posts so far, from the group ( $i_{p-}$ ),
- 740 • from *InfBlogSoc*:
  - to *StdBlog* – lose the post and comments influence as well as high cooperativeness ( $i_{p-c-}$ ),
  - to *StdUsr* – lose the post and comments influence ( $i_{p-}$ ),
  - to *LoAct* – lose the influence of post and their number of posts is below average, so they do not write too many of them ( $i_{p-c-a_{p-}}$ ),
- 745 • from *InfUsrSel*:

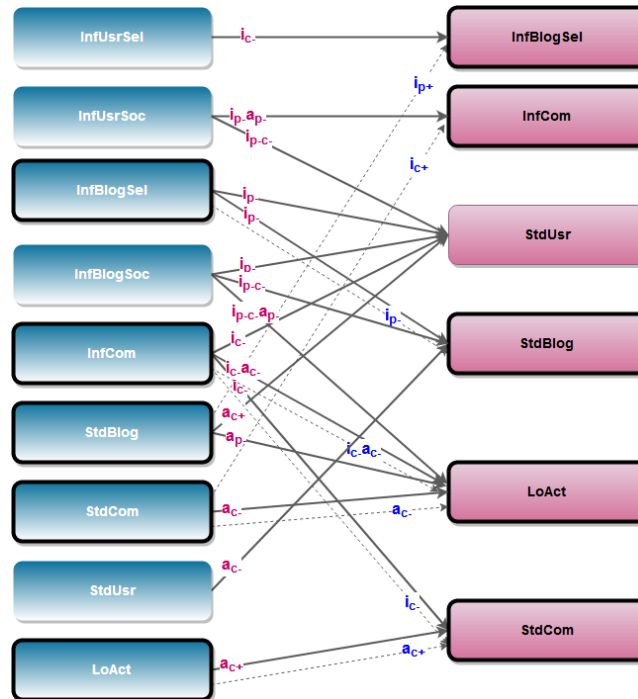


Figure 8: Frequent transitions between roles (blocks with blue background – input roles, blocks with pink background – output roles, frequent transitions for *Salon24* – solid lines, frequent transitions for *The Huffington Post* – dotted line, thick borders of blocks – roles which appear on both portals, thin borders of blocks – roles which appear only in *Salon24*), indexes on lines represent characteristics of the transitions – red for *Salon24*, violet for *The Huffington Post*

- to *InfBlogSel* – lose the comment influence, because of the limitation of the written comments or the limitation of reactions to them ( $i_{c-}$ ),

- from *InfUsrSoc*:

- to *InfCom* – lose the post influence of the post, retains the comment influence, most likely it is related with the lack of written posts and limiting themselves to writing comments ( $i_{p-a_p-}$ ),
- to *StdUsr* – lose post and comment influence, but still write above average numbers of posts and comments, decrease of influence may be caused by the change in the groups' composition ( $i_{p-c-}$ ),

- from *StdCom*:

- to *LoAct* – the numbers of comments decrease from above average to below average ( $a_{c-}$ ),

- from *InfCom*:

- to *StdCom* – comments have ceased to be influential, because of either changes in the composition of the groups or changes on topics commented on by the user – new topics do not provoke further discussions/more comments ( $i_{c-}$ ),
- to *StdUsr* – comments have stopped being influential, their posts and comments are above average numbers ( $i_{c-}$ ),
- to *LoAct* – comments have ceased to be influential, numbers of their posts and comments are below average ( $i_{c-a_c-}$ ),

- from *StdUsr*:

Table 4: Percentage of transitions from a given role to an other role *Salon24*: (1): *StdBlog*, (2): *InfBlogSel*, (3): *InfBlogSoc*, (4): *InfUstrSel*, (5): *InfUstrSoc*, (6): *StdCom*, (7): *InfCom*, (8): *StdUstr*, (9): *LoAct*, (in groups): percent of roles of users which remain in groups in consecutive time slots

from role	to role									(in groups)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	
(1)	61	4	1	0.8	4	1	2	<b>13</b>	<b>15</b>	60.17
(2)	<b>14</b>	47	4	9	2	2	2	<b>10</b>	9	77.20
(3)	<b>11</b>	8	25	2	9	8	6	<b>19</b>	<b>12</b>	95.01
(4)	5	<b>15</b>	1	49	6	2	7	9	4	90.59
(5)	3	2	7	7	44	6	<b>15</b>	<b>10</b>	5	95.01
(6)	0.8	0.4	0.6	0.2	5	17	8	6	<b>23</b>	85.63
(7)	4	1	1	2	4	<b>17</b>	40	<b>13</b>	<b>17</b>	81.60
(8)	<b>12</b>	3	2	2	2	8	8	58	6	87.33
(9)	5	0.7	0.4	0.2	0.2	<b>10</b>	3	2	79	55.44

– to *StdBlog* – their numbers of comments in the group decrease to below the average, as consequences of stopping writing numerous comments or commented users left the group ( $a_{c-}$ ),

• from *LoAct*:

– to *StdCom* – the number of comments written in the groups increased, before they were below average, now they are above it ( $a_{c+}$ ).

For the *The Huffington Post* portal, frequent transitions (marked in bold in Tab 5) are as follows:

• from *StdBlog*:

– to *InfBlogSel* – their already numerous posts have gained influence, they do not comment too much ( $i_{p+}$ ),

• from *InfBlogSel*:

– to *StdBlog* – they lose post influence, like happens in *Salon24* ( $i_{p-}$ ),

• from *StdCom*:

– to *InfCom* – their numerous, above average, comments become influential – this frequent transition did not occur in *Salon24* ( $i_{c+}$ ),

– to *LoAct* – numbers of comments decrease from above the average to below it – as in *Salon24* ( $a_{c-}$ ),

• from *InfCom*:

– *StdCom* – comments ceased to be influential, like in *Salon24* ( $i_{c-}$ ),

– *LoAct* – comments ceased to be influential, his posts and comments are few – below average – as in *Salon24* ( $i_{c-a_{c-}}$ ),

• from *LoAct*:

– to *StdCom* – the number of comments in the group increased, before they were below the average, now they are above it, like in *Salon24* ( $a_{c+}$ ).

Most frequent transitions between roles (omitting the most frequent transitions of maintaining the previous roles) are associated with the demotion of the importance of the role, some are also associated with an increase and decrease of the level of activities. Transitions between influential competitive and cooperative roles are not frequent, this is due to the necessity of introducing fundamental changes in the functioning of a given blogger, so that such a change would be possible. Looking at Tables 3, 4 and 5, one can see that competitive influential roles are characterized by greater stability than cooperative roles, which can lose influence much easier.

Table 5: Percentage of transitions from a given role to an other role *The Huffington Post*: (1): *StdBlog*, (2): *InfBlogSel*, (3): *StdCom*, (4): *InfCom*, (5): *LoAct*, (in groups) percent of roles of users which remain in groups in consecutive time slots

from role	to role					(in groups)
	(1)	(2)	(3)	(4)	(5)	
(1)	75	<b>25</b>	0	0	1	55.53
(2)	<b>39</b>	59	0	0	2	63.56
(3)	0	0	73	<b>10</b>	<b>18</b>	48.70
(4)	0	0	<b>29</b>	51	<b>20</b>	55.30
(5)	0	0	<b>13</b>	6	81	39.48

This clearly shows that activity is an important factor influencing the position that an individual actor performs in a group. Influential roles require high activity, its decline leads to the loss of a position in the group. Users with cooperative roles (*InfBlogSoc* or *InfCom*) are the ones which the least frequently maintain significant roles after transitions. This is due to the fact that occupying the role of an influential blogger, focused not only on engaging in activities aimed at building their own position in the group, but also oriented towards maintaining relationships within the group, requires high activity. Posting and commenting on other users' posts may change over time and they lose recognition when their blog activity is reduced. As a result, they lose the position of an influential blogger and become standard bloggers.

Hence, bloggers, whose influential positions are based on self-directed activity are less likely to lose their position in groups over time. They are focused on maintaining their influential role, which keeps them highly engaged with a high degree of recognition, which is less engaging because it focuses only on publishing one's own posts and writing comments to these posts. This is confirmed by studies conducted by Agarwal et al. (2008); Akritidis et al. (2009); Cai and Chen (2012), who analyzed the interrelationship between bloggers activity and the degree of influence of their position in a group.

In *The Huffington Post*, significant part of users writing posts do it with lower regularity than users in *Salon24*. Given that several important roles (such as *InfUsrSel*, or *InfUsrSoc*) are not identified in *The Huffington Post*, the users of that portal can only increase their importance by changing roles to *InfBlogSel*, but it would require a very fundamental change of their user behavior and much more active and successful post writing (Tab. 5). Taking into consideration specific characteristics of *The Huffington Post*, such a change of role would be very difficult to achieve.

## 6.7. Summary of experiments

The results described in the paper show that individuals tend to maintain their position more often than change it. This indicates that actors strongly identify themselves with roles in online groups, and, similarly as in real life. Assigned roles are more likely to lose influence, while they rarely gain the more influential ones. It is a consequence of the fact, that users usually follow the adapted behavior patterns. The demotion of the role is frequently a consequence of reduction of activities performed on the portal. In large groups, it is harder to become influential, as can be seen in both *Salon24* and *The Huffington Post*. It is easier to be influential and maintain users' influence by consistently building one's own position, rather than disseminate comments in the context of other users' posts.

We identified the frequent transitions between roles on both portals (Fig. 8). One can see similar trends when it comes to frequent user transitions between roles on both sites (taking into account the fact that certain roles, due to the nature of this portal, do not appear on the *The Huffington Post*, which has influenced the limited number of transitions occurring there). This confirms the universality of the adopted definitions of roles.

## 7. Conclusion

In this paper, an analysis of user behavior in social groups in the context of played roles was conducted. A set of roles based on the concept of three dimensions (activity, influence and cooperativeness/competition) which characterize different important features of user behavior justified by sociological approaches and proposed by authors was introduced. Taking advantage on proposed definitions of roles, a model of social system, which defines a set of notions used to describe conducted experiments, was built. The model has a character balanced between generality and detailedness, is flexible and may be extended in the future containing with new elements (e.g., by text and sentiment



analysis, topic modeling, or specific features of other social portals). After performing numerous experiments, several conclusions concerning the impact of group size on the chance of gaining influential roles, or chances of keeping or being promoted to influential roles, were drawn.

A distinction is made between influential roles that are more stable and less stable. Stable roles are the roles which are frequently maintain by users is the subsequent time periods. The most frequent transition from users having each kind of role in the given time periods to the roles these users have in the next time periods were identified. The most frequent reasons of changes of particular roles were identified, they are usually related either to the loss of influence or change (increase or decrease) of activity. We distinguish the most influential roles (*Influential User Social, Influential User Selfish, Influential Blogger Social, Influential Blogger Selfish*), influential but not as influential as these from this first group (*Influential Commentator*), and other roles which are not considered as influential. Translations of users from influential roles to other roles mean that their influence was decreased.

The users most often preserve their roles, but when they change, the demotion of roles occurs more often than its promotion. The most numerous in groups are no active users which do not try to increase their significance in time. The greatest effort needs to be put into preserving influential social (cooperative) roles, which are the most desirable for forming groups. This difficulty comes from the fact, that they are engaged not only in developing their own position but also in supporting the development of significance of other users within the group.

To obtain important roles, it is easier to acquire a role with a high degree of cooperativeness, but such a role is easier to loose than a non-cooperative one. This is why it is most often easier for a user with not important role to gain an important role by obtaining a cooperative version of the important role, but when the importance is achieved, the highest chance to maintain it results in a transition to non-cooperative role which is more stable.

The collected data shows that the dynamics of moving to higher positions is much more difficult than falling to lower positions. This shows that the building up and maintaining an influential role, that requires high recognition, activity and the creation of influential posts and/or comments is related to the hierarchical nature of the leadership structure (Thomas et al. (2013)). Leadership, both focused on building ones own position and being oriented towards social relationships, demands that the actor be active and maintain a high level of recognition.

Conclusions from the conducted analyses may help to better choose activities aimed at gaining an influential position in the community while limiting the costs associated with being up to date with events reflected on a given portal and performing activities significantly affecting other users.

As part of the further work, we want to focus primarily on a more comprehensive analysis of the life cycles of individual user categories, taking into account the different roles they perform either at the level of the whole community, or within the individual groups to which they belong. Aiming at it we intend to introduce modified definitions of roles taking into account also the stability of individual behaviors and his/hers more distant history. We intend to use such representation also for predicting future roles of users, focusing primarily on consideration of a situations where a user strengthens his position, maintains it or loses it. Another direction of further work is the analysis of various other social media, eg. Twitter, using the set of roles presented in the paper, with some possible modifications.

## References

- Abnar, A., Takaffoli, M., Rabbany, R., Zaïane, O.R., 2014. Ssrn: Structural social role mining for dynamic social networks, in: 2014 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM 2014), pp. 289–296. doi:[10.1109/ASONAM.2014.6921599](https://doi.org/10.1109/ASONAM.2014.6921599)
- Agarwal, N., Liu, H., 2009. Modeling and Data Mining in Blogosphere. Morgan & Claypool Publishers.
- Agarwal, N., Liu, H., Tang, L., Yu, P., 2012. Modeling blogger influence in a community. Social Network Analysis and Mining 2, 139–162. doi:[10.1007/s13278-011-0039-3](https://doi.org/10.1007/s13278-011-0039-3)
- Agarwal, N., Liu, H., Tang, L., Yu, P.S., 2008. Identifying the influential bloggers in a community, in: Proc. of the 2008 Int. Conf. on Web Search and Data Mining, ACM, NY, USA. pp. 207–218. doi:[10.1145/1341531.1341559](https://doi.org/10.1145/1341531.1341559)
- Aggarwal, C., Subbian, K., 2014. Evolutionary network analysis: A survey. ACM Comput. Surv. 47, 10:1–10:36. doi:[10.1145/2601412](https://doi.org/10.1145/2601412)
- Agrawal, R., Rajagopalan, S., Srikant, R., Xu, Y., 2003. Mining newsgroups using networks arising from social behavior, in: Proceedings of the 12th International Conference on World Wide Web, Association for Computing Machinery, New York, NY, USA. p. 529–535. doi:[10.1145/775152.775227](https://doi.org/10.1145/775152.775227)
- Ahmed, N.K., Rossi, R.A., Willke, T.L., Zhou, R., 2016. Revisiting role discovery in networks: From node to edge roles. CoRR abs/1610.00844.

- 880 Akritidis, L., Katsaros, D., Bozaris, P., 2009. Identifying influential bloggers: Time does matter, in: 2009 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology, pp. 76–83. doi:[10.1109/WI-IAT.2009.18](https://doi.org/10.1109/WI-IAT.2009.18)
- Anderson, C., 2006. *The Long Tail: Why the Future of Business Is Selling Less of More*. Hyperion.
- Arockiasamy, A., Gionis, A., Tatti, N., 2016. A combinatorial approach to role discovery, in: 2016 IEEE 16th International Conference on Data Mining (ICDM), pp. 787–792. doi:[10.1109/ICDM.2016.0091](https://doi.org/10.1109/ICDM.2016.0091)
- 885 Azaouzi, M., Rhouma, D., Ben Romdhane, L., 2019. Community detection in large-scale social networks: state-of-the-art and future directions. *Social Network Analysis and Mining* 9. doi:[10.1007/s13278-019-0566-x](https://doi.org/10.1007/s13278-019-0566-x)
- Benamar, L., Balagué, C., Ghassany, M., 2017. The identification and influence of social roles in a social media product community. *Journal of Computer-Mediated Communication* 22, 337–362. doi:[10.1111/jcc4.12195](https://doi.org/10.1111/jcc4.12195)
- Benevenuto, F., Rodrigues, T., Cha, M., Almeida, V., 2009. Characterizing user behavior in online social networks, in: Proceedings of the 9th ACM SIGCOMM Conference on Internet Measurement, ACM, New York, NY, USA. pp. 49–62. doi:[10.1145/1644893.1644900](https://doi.org/10.1145/1644893.1644900)
- 890 Berry, J., Keller, E., 2010. *The Influentials: One American in Ten Tells the Other Nine How to Vote, Where to Eat, and What to Buy*. The Free Press.
- Biddle, B.J., 1986. Recent development in role theory. *Annual Review of Sociology* 12, 67–92.
- Bion, W.R., 2003. *Experiences in Groups*. Routledge, New York.
- 895 Blondel, V.D., Guillaume, J.L., Lambiotte, R., Lefebvre, E., 2008. Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment* 2008, P10008. doi:[10.1088/1742-5468/2008/10/p10008](https://doi.org/10.1088/1742-5468/2008/10/p10008)
- Borgatti, S.P., Everett, M.G., 2000. Models of core/periphery structures. *Social Networks* 21, 375 – 395. doi:[10.1016/S0378-8733\(99\)00019-2](https://doi.org/10.1016/S0378-8733(99)00019-2)
- Cai, Y., Chen, Y., 2012. Mining influential bloggers: From general to domain specific, from explicit to implicit. *Int. J. Know.-Based Intell. Eng. Syst.* 16, 223–233. doi:[10.3233/KES-2012-00245](https://doi.org/10.3233/KES-2012-00245)
- 900 Chang, S.E., Liu, A.Y., Shen, W.C., 2017. User trust in social networking services: A comparison of facebook and linkedin. *Computers in Human Behavior* 69, 207 – 217. doi:<https://doi.org/10.1016/j.chb.2016.12.013>
- Coscia, M., Rossetti, G., Giannotti, F., Pedreschi, D., 2012. Demon: A local-first discovery method for overlapping communities, in: KDD'12 - 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 615–623. doi:[10.1145/2339530.2339630](https://doi.org/10.1145/2339530.2339630)
- 905 Coscia, M., Rossetti, G., Giannotti, F., Pedreschi, D., 2014. Uncovering hierarchical and overlapping communities with a local-first approach. *ACM Trans. Knowl. Discov. Data* 9, 6:1–6:27. doi:[10.1145/2629511](https://doi.org/10.1145/2629511)
- Domingos, P., Richardson, M., 2001. Mining the network value of customers, in: Proceedings of the Seventh ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, ACM, New York, NY, USA. pp. 57–66. doi:[10.1145/502512.502525](https://doi.org/10.1145/502512.502525)
- Doreian, P., Batagelj, V., Ferligoj, A., 2004. *Generalized Blockmodeling. Structural Analysis in the Social Sciences*, Cambridge University Press. doi:[10.1017/CBO9780511584176](https://doi.org/10.1017/CBO9780511584176)
- 910 Eccleston, D., Griseri, L., 2008. How does web 2.0 stretch traditional influencing patterns? *International Journal of Market Research - INT J MARKET RES* 50. doi:[10.2501/S1470785308200055](https://doi.org/10.2501/S1470785308200055)
- Eirinaki, M., Monga, S.P.S., Sundaram, S., 2012. Identification of influential social networkers. *Int. J. Web Based Communities* 8, 136–158. doi:[10.1504/IJWBC.2012.046256](https://doi.org/10.1504/IJWBC.2012.046256)
- 915 Farrell, H., Drezner, D., 2007. The power and politics of blogs. *Public Choice* 134, 15–30. doi:[10.1007/s11127-007-9198-1](https://doi.org/10.1007/s11127-007-9198-1)
- Forestier, M., Stavrianou, A., Velcin, J., Zighed, D., 2012. Roles in social networks: Methodologies and research issues. *Web Intelligence and Agent Systems* 10, 117–133. doi:[10.3233/WIA-2012-0236](https://doi.org/10.3233/WIA-2012-0236)
- Fortunato, S., 2010. Community detection in graphs. *Physics Reports* 486, 75 – 174. doi:[10.1016/j.physrep.2009.11.002](https://doi.org/10.1016/j.physrep.2009.11.002)
- Fu, C., 2019. Tracking user-role evolution via topic modeling in community question answering. *Information Processing & Management* 56, 102075. doi:[10.1016/j.ipm.2019.102075](https://doi.org/10.1016/j.ipm.2019.102075)
- 920 Gill, K.E., 2004. How can we measure the influence of the blogosphere?, in: Proceedings of the WW'04:workshop on the Weblogging Ecosystem: Aggregation, Analysis and Dynamics.
- Gillin, P., 2007. *The New Influencers: A Marketer's Guide to the New Social Media*. Quill Driver Books, Sanger, CA.
- Girvan, M., Newman, M.E.J., 2002. Community structure in social and biological networks. *Proceedings of the National Academy of Sciences* 99, 7821–7826. doi:[10.1073/pnas.122653799](https://doi.org/10.1073/pnas.122653799)
- 925 Gleave, E., Welsch, H.T., Lento, T.M., Smith, M.A., 2009. A conceptual and operational definition of 'social role' in online community, in: 2009 42nd Hawaii International Conference on System Sciences, pp. 1–11. doi:[10.1109/HICSS.2009.6](https://doi.org/10.1109/HICSS.2009.6)
- Gliwa, B., Kozlak, J., Zygmunt, A., Cetnarowicz, K., 2012a. Models of social groups in blogosphere based on information about comment addressees and sentiments, in: Proceedings of the 4th International Conference on Social Informatics, Springer-Verlag, Berlin, Heidelberg. pp. 475–488. doi:[10.1007/978-3-642-35386-4\\_35](https://doi.org/10.1007/978-3-642-35386-4_35)
- 930 Gliwa, B., Kozlak, J., Zygmunt, A., Demazeau, Y., 2016. Combining agent-based and social network analysis approaches to recognition of role influence in social media, in: Advances in Practical Applications of Scalable Multi-agent Systems. The PAAMS Collection - 14th International Conference, PAAMS 2016, Sevilla, Spain, June 1-3, 2016, Proceedings, pp. 109–120. doi:[10.1007/978-3-319-39324-7\\_10](https://doi.org/10.1007/978-3-319-39324-7_10)
- Gliwa, B., Saganowski, S., Zygmunt, A., Brodka, P., Kazienko, P., Kozak, J., 2012b. Identification of group changes in blogosphere, in: Proceedings of the 2012 International Conference on Advances in Social Networks Analysis and Mining (ASONAM 2012), IEEE Computer Society, Washington, DC, USA. pp. 1201–1206. doi:[10.1109/ASONAM.2012.207](https://doi.org/10.1109/ASONAM.2012.207)
- 935 Gliwa, B., Zygmunt, A., Kozlak, J., 2013. Analysis of roles and groups in blogosphere, in: Burduk, R., Jackowski, K., Kurzynski, M., Wozniak, M., Zolnierek, A. (Eds.), Proceedings of the 8th International Conference on Computer Recognition Systems CORES 2013. Springer International Publishing. volume 226 of *Advances in Intelligent Systems and Computing*, pp. 299–308. doi:[10.1007/978-3-319-00969-8\\_29](https://doi.org/10.1007/978-3-319-00969-8_29)
- 940 Golder, S., Donath, J., 2004. Social roles in electronic communities. *Internet Research* 5.
- Greene, D., Doyle, D., Cunningham, P., 2010. Tracking the evolution of communities in dynamic social networks, in: Proceedings of the 2010 International Conference on Advances in Social Networks Analysis and Mining, IEEE Computer Society, Washington, DC, USA. pp. 176–183. doi:[10.1109/ASONAM.2010.17](https://doi.org/10.1109/ASONAM.2010.17)
- Gregory, S., 2009. Finding overlapping communities using disjoint community detection algorithms, in: *Complex Networks - Results of the 2009*

- 945 International Workshop on Complex Networks, CompleNet 2009, Catania, Italy, pp. 47–61. doi:[10.1007/978-3-642-01206-8\\_5](https://doi.org/10.1007/978-3-642-01206-8_5)
- Gritsenko, V., 2016. Interaction on online forums and group communication: A case study of an it support community. *Procedia - Social and Behavioral Sciences* 236, 14 – 24. doi:[10.1016/j.sbspro.2016.12.004](https://doi.org/10.1016/j.sbspro.2016.12.004) international Conference on Communication in Multicultural Society, CMSC 2015, 6-8 December 2015, Moscow, Russian Federation.
- Gruhl, D., Guha, R., Liben-Nowell, D., Tomkins, A., 2004. Information diffusion through blogspace, in: Proceedings of the 13th International Conference on World Wide Web, ACM, New York, NY, USA. pp. 491–501. doi:[10.1145/988672.988739](https://doi.org/10.1145/988672.988739)
- 950 Guimera, R., Amaral, L.A.N., 2005. Functional cartography of complex metabolic networks. *Nature* 433, 895–900.
- Hanneman, R.A., Riddle, M., 2005. Introduction to social network methods. University of California, Riverside, Riverside, CA.
- Hare, A.P., 1994. Types of roles in small groups: A bit of history and a current perspective. *Small Group Research* 25, 433–448. doi:[10.1177/1046496494253005](https://doi.org/10.1177/1046496494253005)
- 955 Heidemann, J., Klier, M., Probst, F., 2012. Online social networks: A survey of a global phenomenon. *Computer Networks* 56, 3866 – 3878. doi:[10.1016/j.comnet.2012.08.009](https://doi.org/10.1016/j.comnet.2012.08.009) the WEB we live in.
- Henderson, K., Gallagher, B., Eliassi-Rad, T., Tong, H., Basu, S., Akoglu, L., Koutra, D., Faloutsos, C., Li, L., 2012. Rolx: structural role extraction & mining in large graphs., in: Yang, Q., Agarwal, D., Pei, J. (Eds.), KDD, ACM. pp. 1231–1239.
- Herrmann, T., Jahnke, I., Loser, K.U., 2004. The role concept as a basis for designing community systems., in: COOP, pp. 163–178.
- 960 Hirvonen, P., 2016. Positioning theory and small-group interaction: Social and task positioning in the context of joint decision-making. *SAGE Open* 6. doi:[10.1177/2158244016655584](https://doi.org/10.1177/2158244016655584)
- Homans, G.C., 1950. *The Human Group*. Harcourt, Brace and Company, New York.
- Hozhyi, V., Lamiroy, B., 2017. Clustering of users in social networks by their activity, in: 2017 IEEE First Ukraine Conference on Electrical and Computer Engineering (UKRCON), pp. 1072–1077. doi:[10.1109/UKRCON.2017.8100415](https://doi.org/10.1109/UKRCON.2017.8100415)
- 965 Hromic, H., Hayes, C., 2018. Characterising and evaluating online communities from live microblogging user interactions, pp. 21–24. doi:[10.1109/ASONAM.2018.8508392](https://doi.org/10.1109/ASONAM.2018.8508392)
- Johnson, S., Safadi, H., Faraj, S., 2015. The emergence of online community leadership. *Information Systems Research* 26, 165–187. doi:[10.1287/isre.2014.0562](https://doi.org/10.1287/isre.2014.0562)
- Junquero-Trabado, V., Dominguez-Sal, D., 2012. Building a role search engine for social media. WWW'12 - Proceedings of the 21st Annual Conference on World Wide Web Companion doi:[10.1145/2187980.2188240](https://doi.org/10.1145/2187980.2188240)
- 970 Katona, Z., Zubcsek, P.P., Sarvary, M., 2011. Network effects and personal influences: The diffusion of an online social network. *Journal of Marketing Research* 48, 425–443. doi:[10.1509/jmkr.48.3.425](https://doi.org/10.1509/jmkr.48.3.425)
- Kempe, D., Kleinberg, J., Tardos, E., 2003. Maximizing the spread of influence through a social network, in: Proceedings of the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, ACM, New York, NY, USA. pp. 137–146. doi:[10.1145/956750.956769](https://doi.org/10.1145/956750.956769)
- 975 Kenneth, C., Yang, C., 2011. The effects of social influence on blog advertising use. *Journal of Intercultural Communication Studies* 2, 131–147.
- Khan, H.U., Daud, A., Ishfaq, U., Amjad, T., Aljohani, N., Abbasi, R.A., Alowibdi, J.S., 2017. Modelling to identify influential bloggers in the blogosphere: A survey. *Computers in Human Behavior* 68, 64 – 82. doi:[10.1016/j.chb.2016.11.012](https://doi.org/10.1016/j.chb.2016.11.012)
- Kim, E.S., Han, S.S., 2009. An analytical way to find influencers on social networks and validate their effects in disseminating social games, in: Proceedings of the 2009 International Conference on Advances in Social Network Analysis and Mining, IEEE Computer Society, Washington, DC, USA. pp. 41–46. doi:[10.1109/ASONAM.2009.59](https://doi.org/10.1109/ASONAM.2009.59)
- 980 Kimura, M., Saito, K., Nakano, R., Motoda, H., 2010. Extracting influential nodes on a social network for information diffusion. *Data Min. Knowl. Discov.* 20, 70–97. doi:[10.1007/s10618-009-0150-5](https://doi.org/10.1007/s10618-009-0150-5)
- Koźlak, J., Zygmunt, A., Gliwa, B., Rudek, K., 2018. Dynamics of social roles in the context of group evolution in the blogosphere, in: 2018 5th International Conference on Behavioral, Economic, and Socio-Cultural Computing (BESC), pp. 179–184. doi:[10.1109/BESC.2018.8697251](https://doi.org/10.1109/BESC.2018.8697251)
- 985 Kumar, S., Morstatter, F., Liu, H., 2013. *Twitter Data Analytics*. Springer Publishing Company, Incorporated.
- Labatut, V., Dugué, N., Perez, A., 2014. Identifying the community roles of social capitalists in the Twitter network, in: 2014 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM 2014), pp. 371–374. doi:[10.1109/ASONAM.2014.6921612](https://doi.org/10.1109/ASONAM.2014.6921612)
- 990 Lancichinetti, A., Radicchi, F., Ramasco, J.J., Fortunato, S., 2011. Finding statistically significant communities in networks. *PLoS One* 6. doi:[10.1371/journal.pone.0018961](https://doi.org/10.1371/journal.pone.0018961)
- Libai, B., Bolton, R., Bügel, M., de ruyter, k., Götz, O., Risselada, H., T. Stephen, A., 2010. Customer-to-customer interactions: Broadening the scope of word of mouth research. *Journal of Service Research - J SERV RES* 13, 267–282. doi:[10.1177/1094670510375600](https://doi.org/10.1177/1094670510375600)
- 995 Long, Y., Li, V.O.K., Niu, G., Lu, Z., 2015. The impacts of network structure on user activity level in online social networks, in: 2015 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT), pp. 115–122. doi:[10.1109/WI-IAT.2015.201](https://doi.org/10.1109/WI-IAT.2015.201)
- Makagon, M., Mccowan, B., Mench, J., 2012. How can social network analysis contribute to social behavior research in applied ethology? *Applied Animal Behaviour Science* 138, 152–161. doi:[10.1016/j.applanim.2012.02.003](https://doi.org/10.1016/j.applanim.2012.02.003)
- 1000 McNeill, A.R., Briggs, P., 2014. Understanding twitter influence in the health domain: A social-psychological contribution, in: Proceedings of the 23rd International Conference on World Wide Web, ACM, New York, NY, USA. pp. 673–678. doi:[10.1145/2567948.2579280](https://doi.org/10.1145/2567948.2579280)
- Nath, K., Roy, S., 2019. Detecting intrinsic communities in evolving networks. *Social Network Analysis and Mining* 9, 13. doi:[10.1007/s13278-019-0552-3](https://doi.org/10.1007/s13278-019-0552-3)
- Newman, M.E.J., Park, J., 2003. Why social networks are different from other types of networks. *Phys. Rev. E* 68, 036122.
- 1005 Nosenzo, D., Quercia, S., Sefton, M., 2015. Cooperation in small groups: the effect of group size. *Experimental Economics* 18, 4–14. doi:[10.1007/s10683-013-9382-8](https://doi.org/10.1007/s10683-013-9382-8)
- Palla, G., Ábel, D., Farkas, I.J., Pollner, P., Derényi, I., Vicsek, T., 2008. k-Clique Percolation and Clustering. Springer Berlin Heidelberg, Berlin, Heidelberg. pp. 369–408. doi:[10.1007/978-3-540-69395-6\\_9](https://doi.org/10.1007/978-3-540-69395-6_9)
- Palla, G., Barabasi, A.L., Vicsek, T., 2007. Quantifying social group evolution. *Nature* 446, 664–667.
- Palla, G., Der[acute]nyi, I., Farkas, I., Vicsek, T., 2005. Uncovering the overlapping community structure of complex networks in nature and

- society. *Nature* 435, 814–818. doi:[10.1038/nature03607](https://doi.org/10.1038/nature03607).
- Parthasarathy, S., Ruan, Y., Satuluri, V., 2011. Community Discovery in Social Networks: Applications, Methods and Emerging Trends. pp. 79–113. doi:[10.1007/978-1-4419-8462-3\\_4](https://doi.org/10.1007/978-1-4419-8462-3_4)
- Patel, H., Pettitt, M., Wilson, J., 2011. Factors of collaborative working: A framework for a collaboration model. *Applied ergonomics* 43, 1–26. doi:[10.1016/j.apergo.2011.04.009](https://doi.org/10.1016/j.apergo.2011.04.009)
- Pei, Y., Fletcher, G., Pechenizkiy, M., 2019. Joint role and community detection in networks via  $l_{2,1}$  norm regularized nonnegative matrix tri-factorization, in: *Proceedings of the 2019 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*.
- Pei, Y., Zhang, J., Fletcher, G., Pechenizkiy, M., 2018. DyNMF: Role analytics in dynamic social networks, pp. 3818–3824. doi:[10.24963/ijcai.2018/531](https://doi.org/10.24963/ijcai.2018/531)
- Pluciennik, J., 2013. Measurement of activity in social media – a challenge for contemporary marketing research. *Marketing of Scientific and Research Organisations* 8. doi:[10.14611/minib.08.02.2013.12](https://doi.org/10.14611/minib.08.02.2013.12)
- Rossi, R., Ahmed, N., 2014. Role discovery in networks. *IEEE Transactions on Knowledge and Data Engineering* 27. doi:[10.1109/TKDE.2014.2349913](https://doi.org/10.1109/TKDE.2014.2349913)
- Rossi, R., Gallagher, B., Neville, J., Henderson, K., 2012. Role-dynamics: Fast mining of large dynamic networks, in: *Proceedings of the 21st International Conference on World Wide Web*, ACM, New York, NY, USA. pp. 997–1006. doi:[10.1145/2187980.2188234](https://doi.org/10.1145/2187980.2188234)
- Rossi, R.A., Gallagher, B., Neville, J., Henderson, K., 2013. Modeling dynamic behavior in large evolving graphs, in: *Proceedings of the Sixth ACM International Conference on Web Search and Data Mining*, ACM, New York, NY, USA. pp. 667–676. doi:[10.1145/2433396.2433479](https://doi.org/10.1145/2433396.2433479)
- Rullani, F., Haefliger, S., 2013. The periphery on stage: The intra-organizational dynamics in online communities of creation. *Research Policy* 42, 941–953. doi:[10.1016/j.respol.2012.10](https://doi.org/10.1016/j.respol.2012.10).
- Scripps, J., Tan, P.N., Esfahanian, A.H., 2007. Node roles and community structure in networks, in: *Proceedings of the 9th WebKDD and 1st SNA-KDD 2007 Workshop on Web Mining and Social Network Analysis*, ACM, New York, NY, USA. pp. 26–35. doi:[10.1145/1348549.1348553](https://doi.org/10.1145/1348549.1348553)
- Song, X., Chi, Y., Hino, K., Tseng, B., 2007. Identifying opinion leaders in the blogosphere, in: *Proceedings of the Sixteenth ACM Conference on Conference on Information and Knowledge Management*, ACM, New York, NY, USA. pp. 971–974. doi:[10.1145/1321440.1321588](https://doi.org/10.1145/1321440.1321588)
- Spiliopoulou, M., 2011. Evolution in social networks: A survey., in: Aggarwal, C.C. (Ed.), *Social Network Data Analytics*. Springer, pp. 149–175.
- Stephen, A., Dover, Y., Goldenberg, J., 2010. You snooze you loose: Comparing the roles of high activity and connectivity in information dissemination over online social networks, in: Dahl, D.W., Johar, G.V., van Osselaer, S.M. (Eds.), *NA - Advances in Consumer Research Volume 38*. Duluth, MN : Association for Consumer Research.
- Takaffoli, M., Fagnan, J., Sangi, F., Zaïane, O.R., 2011. Tracking changes in dynamic information networks, in: *2011 International Conference on Computational Aspects of Social Networks (CASoN)*, pp. 94–101. doi:[10.1109/CASON.2011.6085925](https://doi.org/10.1109/CASON.2011.6085925)
- Teh, P.L., Pak, I., Rayson, P., Piao, S., 2015. Exploring fine-grained sentiment values in online product reviews, in: *2015 IEEE Conference on Open Systems (ICOS)*, pp. 114–118. doi:[10.1109/ICOS.2015.7377288](https://doi.org/10.1109/ICOS.2015.7377288)
- Tennakoon, G., Nayak, R., 2019. FCMiner: mining functional communities in social networks. *Social Network Analysis and Mining* 9. doi:[10.1007/s13278-019-0565-y](https://doi.org/10.1007/s13278-019-0565-y)
- Thomas, G., Martin, R., Riggio, R., 2013. Leading groups: Leadership as a group process. *Group Processes & Intergroup Relations* 16, 3–16. doi:[10.1177/1368430212462497](https://doi.org/10.1177/1368430212462497)
- Turner, R., 1968. Social roles: Sociological aspects, in: Skills, D.L., Merton, R.K. (Eds.), *International Encyclopedia of the Social Sciences*. Macmillan, New York.
- Turner, R., 1978. The role and the person. *American Journal of Sociology* 84, 1–23.
- Vaca Ruiz, C., Aiello, L.M., Jaimés, A., 2014. Modeling dynamics of attention in social media with user efficiency. *EPJ Data Science* 3, 5. doi:[10.1140/epjds30](https://doi.org/10.1140/epjds30)
- Varlamis, I., Eiriniaki, M., Louta, M., 2013. Application of Social Network Metrics to a Trust-Aware Collaborative Model for Generating Personalized User Recommendations. Springer Vienna, Vienna. pp. 49–74. doi:[10.1007/978-3-7091-1346-2\\_3](https://doi.org/10.1007/978-3-7091-1346-2_3)
- Wasserman, S., Faust, K., 1994. *Social network analysis: Methods and applications*. volume 8. Cambridge university press.
- Welser, H.T., Gleave, E., Fisher, D., Smith, M.A., 2007. Visualizing the signatures of social roles in online discussion groups. *Journal of Social Structure* 8.
- Weng, J., Lim, E.P., Jiang, J., He, Q., 2010. Twitterrank: Finding topic-sensitive influential twitterers, in: *Proceedings of the Third ACM International Conference on Web Search and Data Mining*, ACM, New York, NY, USA. pp. 261–270. doi:[10.1145/1718487.1718520](https://doi.org/10.1145/1718487.1718520)
- Zafarani, R., Abbasi, M.A., Liu, H., 2014. *Social media mining: an introduction*. Cambridge University Press.
- Zhu, Y., Zhong, E., Pan, S.J., Wang, X., Zhou, M., Yang, Q., 2013. Predicting user activity level in social networks, in: *Proceedings of the 22Nd ACM International Conference on Information & Knowledge Management*, ACM, New York, NY, USA. pp. 159–168. doi:[10.1145/2505515.2505518](https://doi.org/10.1145/2505515.2505518)
- Zygmunt, A., 2018. Role identification of social networkers, in: *Encyclopedia of Social Network Analysis and Mining*, 2nd Edition, pp. 1598–1606. doi:[10.1007/978-1-4939-7131-2\\_247](https://doi.org/10.1007/978-1-4939-7131-2_247)
- Zygmunt, A., Bródka, P., Kazienko, P., Kozlak, J., 2011. Different approaches to groups and key person identification in blogosphere, in: *International Conference on Advances in Social Networks Analysis and Mining, ASONAM 2011, Kaohsiung, Taiwan, 25-27 July 2011*, pp. 593–598. doi:[10.1109/ASONAM.2011.71](https://doi.org/10.1109/ASONAM.2011.71)
- Zygmunt, A., Bródka, P., Kazienko, P., Kozlak, J., 2012. Key person analysis in social communities within the blogosphere. *J. UCS* 18, 577–597. doi:[10.3217/jucs-018-04-0577](https://doi.org/10.3217/jucs-018-04-0577)
- Zygmunt, A., Gliwa, B., 2015. The comparison of users activity on the example of polish and american blogosphere. *Sci. Program.* 2015, 907547:1–907547:11. doi:[10.1155/2015/907547](https://doi.org/10.1155/2015/907547)
- Zygmunt, A., Kozlak, J., Gliwa, B., 2018. Roles in local communities and global position in social media, in: *2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*, pp. 1204–1211. doi:[10.1109/ASONAM.2018.8508258](https://doi.org/10.1109/ASONAM.2018.8508258)