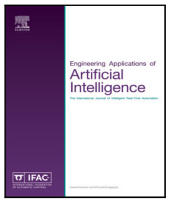
Engineering Applications of Artificial Intelligence 99 (2021) 104129

Contents lists available at ScienceDirect

Engineering Applications of Artificial Intelligence

journal homepage: www.elsevier.com/locate/engappai



Analysing centralities for organisational role inference in online social networks

Rubén Sánchez-Corcuera *, Aritz Bilbao-Jayo, Unai Zulaika, Aitor Almeida

University of Deusto - DeustoTech, Bilbao, Spain
Unibertsitate Etorbidea, 24, 48007 Bilbo, Bizkaia, Spain

ARTICLE INFO

Keywords:
Online social networks
Adversarial information retrieval
Information inference
Graph centralities

ABSTRACT

The intensive use of Online Social Networks (OSN) nowadays has made users expose more information without realising it. Malicious users or marketing agencies are now able to infer information that is not published on OSNs by using data from targets friends to use for their benefit. In this paper, the authors present a generalisable method capable of deducing the roles of employees of an organisation using their Twitter relationships and the features of the graph from their organisation. The authors also conduct an extensive analysis of the node centralities to study their roles in the inference of the different classes proposed. Derived from the experiments and the ablation study conducted to the centralities, the authors conclude that the latent features of the graph along with the directed relationships perform better than previously proposed methods when classifying the role of the employees of an organisation. Additionally, to evaluate the method, the authors also contribute with a new dataset consisting of three directed graphs (one for each organisation) representing the relationships between the employees obtained from Twitter.

1. Introduction

Online social networks (OSN) have become part of everyday life for millions of Internet users. These tools have changed how people interact by providing new communication channels and features to share the content of interest between people all over the world. Social networks have been analysed from different academic disciplines since their inception (Howard, 2008; Steinfield et al., 2008; Garton et al., 1997; Saltz et al., 2004). The content published on them may reveal valuable information that can be used to make decisions or perform actions in real life. For example, emergency teams may know and act on an emergency sooner (Kim and Hastak, 2018) or users may infer the political preference of a group of people who interact on a social network (Bilbao-Jayo and Almeida, 2018). In addition to the analysis of the content published in social networks, many interested individuals, such as marketing companies or malicious users, may want to know private information from users in order to increase their knowledge of them. This phenomenon is called private information inference.

Private information inference has been thoroughly investigated by researchers in the area leading to the publishing of many works (Fire and Puzis, 2016; Altenburger and Ugander, 2018; Chaabane et al., 2012; Chaniotakis et al., 2017; Zhang et al., 2016; Gong and Liu, 2016; Fani et al., 2019; Zarrinkalam et al., 2018; Valverde-Rebaza et al., 2018; Chen et al., 2016). Private information inference is categorised in friend-based or behaviour-based depending if the information used for the inference is extracted from the relationships of the target (friend-based) or from the content the target publishes on the OSN (behaviour-based) (Gong and Liu, 2018).

Behaviour-based methods employ data published on online social networks such as images, text or video. On the one hand, this information can help to infer many data about users since, as has been shown, the way users write or the photos they post on the networks are representative of their personality (Li et al., 2012; Thomas et al., 2010; Chaabane et al., 2012; You et al., 2014). On the other hand, due to the heterogeneous nature of this data, the possibility of generalising this approach among different OSNs is limited.

Instead, friend-based approaches use the relationships between users and public information to extract private information from the targets. One of the properties that made this possible is the homophilia, which describes that users tend to relate to those whom they share most attributes with (McPherson et al., 2001), for example, sex, employment or hobbies. Also, relationships between users are straightforward to model through a graph, which allows access to the centralities of the nodes and extract information about the most relevant and influential users in the network.

Node centralities have been created to identify the most important or influential node in a graph taking into account different characteristics of the network topology to calculate them. For example,

* Corresponding author at: University of Deusto - DeustoTech, Bilbao, Spain.
E-mail addresses: ruben.sanchez@deusto.es (R. Sánchez-Corcuera), aritzbilbao@deusto.es (A. Bilbao-Jayo), unai.zulaika@deusto.es (U. Zulaika), aitor.almeida@deusto.es (A. Almeida).

https://doi.org/10.1016/j.engappai.2020.104129
Received 28 July 2020; Received in revised form 27 October 2020; Accepted 2 December 2020
Available online 30 December 2020

degree centrality (Opsahl et al., 2010) allows to find the nodes which are in the "middle" of the network; instead, betweenness centrality (Newman, 2010) helps to find the nodes that are "bridges" between separated parts of the graphs, thus having into account the whole graph. Thus, mixing the information provided by the different centralities researchers have been able to model behaviours such as, identifying the traffic congestion on the road network (Jayaweera et al., 2017), analyse relationships between students in classrooms (Grunspan et al., 2014), analyse biological networks to find correlations between them (Koschützki et al., 2005) or analyse global structure the air transportation network worldwide (Guimera et al., 2005). For this reason, we have decided to use different centralities in our approach to infer private information seeking to create a model that is both generalisable and effective.

1.1. Objectives & contributions

In this paper, we focus on the inference of hierarchical roles of employees of organisations using data from users on Twitter OSN. Through these experiments, an attempt has been made to highlight the ease with which private data can be extracted from users of social networks. Besides, we believe that the hierarchy of an organisation can enable an attack vector for malicious users by revealing who are the people with higher positions within the organisation. For this reason, through this paper, we wanted to carry out experiments to demonstrate that it is possible to infer the hierarchy of an organisation using only the relationships of their members on Twitter.

To this end, we propose to use data extracted from the relations of those users on the online social network as features for several classification algorithms to obtain better results. We have modelled the user relations as a graph and extracted the centralities of the nodes. Centralities explained in Section 3.2, have been used previously to classify nodes due to the information they offer about them and the network in general (Diallo et al., 2016; Zhan et al., 2017).

We evaluate and test our approach under different settings and datasets and compare the obtained performance against the current best approach in the state-of-the-art, which was proposed by Fire and Puzis (2016). For the experiments, we create a dataset that contains employees with their roles from three different organisations gathered from Twitter OSN. We also focus on evaluating our method in generalisation situation so we can assess the performance of the algorithm when used in other scenarios.

The main contributions of the paper can be summarised as:

- We propose a new set of centralities as features that improve the inference of the position in the organisation of a user present in a social network.
- We present an in-depth analysis of node centralities and the information they provide for the inference of the hierarchy of an organisation.
- We provide a new labelled dataset containing graphs that represent the relationships of members from three different organisations on Twitter.[1]

The rest of the paper is organised as follows: We first provide background information about the OSNs and the adversarial information retrieval done in them in Section 2. After that, in Section 3, we present the dataset that we used for the experiments, how we gathered it and the attributes they have. In Section 4, we describe the methodology we used for the experiments and the results obtained in them. Experiments are followed by Section 5 in which we discuss the results presented in the previous section. Finally, we present our conclusions and suggest some future work in Section 6.

2. Related work

Social networks offer the capacity for sharing information that can be used by other people with good or bad intentions. Adversarial information retrieval has been a security concern in social networks since their inception. Since the lastest trends given in OSNs, such as fake news on political matters (Grinberg et al., 2016; Bovet and Makse, 2016; Bastos and Mercea, 2019) and the rise of social bots (Ferrara et al., 2016; Varol et al., 2017; Davis et al., 2016), the literature in this area has grown considerably. The recovery of adverse information may be carried out using different methods; however, in this paper, we will focus on the inference of private information using the public information available in Online Social Networks.

The proposed approach and analysis implemented in this article is based on the inference of information from public data available in OSNs. As stated by Gong and Liu (2018), the inference attacks conducted in OSNs can be classified into two categories: *friend-based* (Fire and Puzis, 2016; Valverde-Rebaza et al., 2018; Chen et al., 2016) and *behaviour-based* (Zhang et al., 2016; Gong and Liu, 2016; Fani et al., 2019; Zarrinkalam et al., 2018). Friend-based attacks use the relationships and the attributes of friends to infer the desired information. These attacks are based on the premise called *homophily* (McPherson et al., 2001), which states that users are more attracted to similars. Instead, behaviour-based attacks use activity from users and some information from users friends to infer information such as nationality or hobbies.

Regarding currently performed works on information inference on OSNs, we have identified two categories: activity inference and personal attributes inference. Work that aims at inferring activities uses the information publicly available on social networks to deduce which activities are being performed by the target users. In work proposed by Noulas et al. (2013) they propose an activity inference methodology by using data from the OSN Foursquare[2] combined along with data from a telecommunication provider in Spain. Authors propose a classification task in which the telecommunication data has to be associated with the Foursquare semantic labels. Similar work was proposed by Chaniotakis et al. (2017) in which a combination of public data from Twitter and Foursquare was used to infer the activities conducted by Londoners. Authors use a data enrichment method adding information about location and Foursquare activity tagging into tweets produced by the target users to classify them into an activity taxonomy. Although these works are centred in activity inference and mostly behaviour-based, the classification and data gathering methodologies used in them are similar to the ones we proposed.

Works proposed for the inference of attributes in social networks are more common than those of activity inference. Related to attribute inference, many works aim to infer several or a specific attribute from target user employing supervised learning machine learning algorithms with information extracted from the OSN (Zarrinkalam et al., 2018; Zhang et al., 2016; Chen et al., 2016; Mei et al., 2017; Mulders et al., 2019; Gong and Liu, 2016). Many researchers have studied this topic because of the interest that many organisations have and the recent attacks on Online Social Networks. For this reason, many researchers attempted to generalise the attribute inference and proposed models that can classify every possible attribute for a user, for example, AttriInfer. Jia et al. (2017) proposed a model based on modelling the social network as a pairwise Markov Random Field. This approach is based on both behaviour and friend information to decide if a target user will have a specific attribute.

Finally, there are specific works that have conducted hierarchical role inference on OSNs. Chen et al. (2016) published a paper in which they used a *friend-based* approach to infer the roles of Google employees in Google+ OSN. Authors proposed a Naive–Bayes-based model that

[1] https://github.com/rubensancor/RoleMining_Twitter/.

[2] https://foursquare.com/.

uses network metrics such as, Degree Centrality or Cluster Coefficient and neighbour information to infer users roles with promising results. Fire and Puzis (2016) proposed to use scraped information from Facebook accounts of members of organisations to infer the leadership positions at the communities on those organisations using supervised learning classifiers. Unlike these studies, our work incorporates directed graphs and a selected set of centralities that add more information to the machine learning algorithms allowing them to improve results. On the other hand, our work aims to generalise the inference of roles to other organisations not yet seen by the algorithm. Also, regarding the work of Fire & Puzis, our approach includes three categories of employees instead of detecting the leader of the organisation.

3. Research methodology

In this paper, we focused mainly on the inference of the positions of employees in an organisation. To do so, we have conducted experiments to achieve better performance in inferring the role of employees in an organisation by using data from their relationships on Twitter. Furthermore, we have developed an algorithm based on the work done by Fire and Puzis, capable of retrieving employees from the desired organisation on Twitter. Following this, we have developed a dataset composed by the data retrieved from three companies with the algorithm mentioned before.

3.1. Data collection

For this study, we collected data from Twitter accounts belonging to employees from three different companies. The process of data collection is divided into two phases. The first one is to gather users that may be employees of the desired companies; the second one is to consult if the users that are retrieved are real employees of the company and their roles in it by visiting their LinkedIn profiles. A chart summarising the employed methodology is represented in Fig. 1.

To gather the dataset, we followed the methodology employed by Fire and Puzis for their experiments so the results could be comparable. For the first phase, we adapted the algorithm created by Fire and Puzis to make it work on Twitter. This algorithm (Algorithm 1) should have some employees from the desired company defined to start gathering their followers and followees to check if they belong to the company by checking if their description contains a series of keywords. Subsequently, the algorithm continues with this process with the new users it has retrieved. The need for the name of the organisation or one of the defined keywords to be present in the biography of users presents a limitation that can reduce the effectiveness of this method.

Through this method, represented in green in Fig. 1, we gathered data from three different organisations present on Twitter. The first organisation (Organisation A) is a multinational professional services firm with more than 200,000 employees according to their website. The crawler gathered 2988 different Twitter accounts in a week: 2706 employees' accounts, 204 corporate accounts and 78 accounts that were somehow related to the company but were not employees.

The second organisation (Organisation B) is a multinational engineering company that employs more than 350,000 people all over the world. In this case, 1143 accounts were gathered, 807 belonged to the organisation's employees, 163 corporate accounts and 173 that were not related to the organisation.

The third organisation (Organisation C) is a multinational corporation that provides communication technology and services all over the world and according to their website employs 116,000 people. 834 accounts were gathered: 603 from employees, 78 corporate accounts and 116 accounts which were not employees of the company. A summary with the number of employees per organisation and the accuracy achieved in each of them can be seen in Table 1.

Table 1 shows that the algorithm achieved high accuracy rates in the crawl conducted to gather users from Twitter. The accuracy

Algorithm 1 Organisational Mining Algorithm from (Fire and Puzis, 2016) modified for Twitter. The default priority is set in 30.

Input: A set of seed Twitter Usernames (S) of organisation's employees and a set of words related to the target organisation, N

Output: A set of Twitter profiles with twitterid, name, username, biography, followers and followings

Organisational Miner() :

$Q \leftarrow Priority - Queue()$
$\forall_{Username \epsilon S}, Q.Enqueue(Username : 30)$
$Crawled \leftarrow \emptyset$
$NonRelatedUsers \leftarrow 0$
while $(Q \neq \emptyset \ \varsigma \ max(Q.Priority) \neq 1 \ \varsigma \ NonRelatedUsers < 1000)$ **do**
  $Username \leftarrow Q.Dequeue()$
  $Page \leftarrow DownloadTwitterProfileData(Username)$
  $Crawled \leftarrow Crawled.append(Username)$
  if $N\ in\ Page$ **then**
    $Connections \leftarrow ExtractConnFromTwitter()$
    $Connections \leftarrow Connections - Crawled$
    for $(Connection\ \epsilon\ (Connections \cap Q))$ **do**
      $Increasepriority(Connection)$
    end for
    for $(Connection\ \epsilon\ (Connections - Q))$ **do**
      $Q.Enqueue(Connection, Priority : 1)$
    end for
    $CollectedPages.append(Page)$
  else
    $NonRelatedUsers++$
  end if
end while
return $CollectedPages$

was calculated by dividing the number of employees gathered by the number of accounts gathered. As we can see in the table, the accuracy for the first organisation (O1) is higher than the average. The main reason for this is that we used more seeds for the initial step of the crawling algorithm. Also, this accuracy may be influenced by the descriptions the employees use in their online profiles. Furthermore, we also gathered corporate accounts that were not taken into account previously by Fire and Puzis (2016).

After that, we tagged every user in the dataset with its role in the organisation, represented in orange in Fig. 1. Users were classified in three different roles depending on their position on the company from high to low in the hierarchy: executives, managers and employees. This process was done using public data from LinkedIn. During this process, we had to discard 284 profiles because we were unable to decide if they belonged to the organisation since the information was not public or was very ambiguous. Table 2 show the distribution per class in each organisation.

Finally, after preprocessing and cleaning the collected data we transformed the data from each organisation into a directed graph. We employed directed graphs representing the follows of Twitter users; this presents a difference with the graphs analysed by Fire and Puzis (2016). In their work, they use graphs extracted from Facebook relationships where they are bidirectional, so the graphs are undirected. The attributes of the graphs generated for each organisation are represented in Table 1. Furthermore, the graphs extracted from Twitter represent organisations of three different sizes; therefore our approach is tested and trained with data from organisations of different sizes, so the results are representative.

3.2. Feature extraction

In graph theory and network analysis, centralities stratify nodes taking into account different characteristics of their relations with other

Table 1
Summary of the gathered data. Non-related accounts contain corporate accounts and other accounts. The nodes are less than the № of accounts because some profiles were discarded as we were unable to decide if they belonged to the organisation.

Organisation	№ of accounts	№ of employees	№ of non-related accounts	Accuracy	Nodes	Edges
Organisation A	2988	2706	204 + 78	90.5%	2522	36 101
Organisation B	1143	807	163 +173	70.6%	762	3201
Organisation C	834	640	78 + 116	76.7%	585	5482

Table 2
Distribution of roles per organisation.

Organisation	Employees	Managers	Executives
Organisation A	1787 (70.85%)	698 (27.68%)	37 (1.47%)
Organisation B	386 (50.65%)	319 (41.86%)	57 (7.49%)
Organisation C	408 (69.74%)	165 (28.21%)	12 (2.05%)

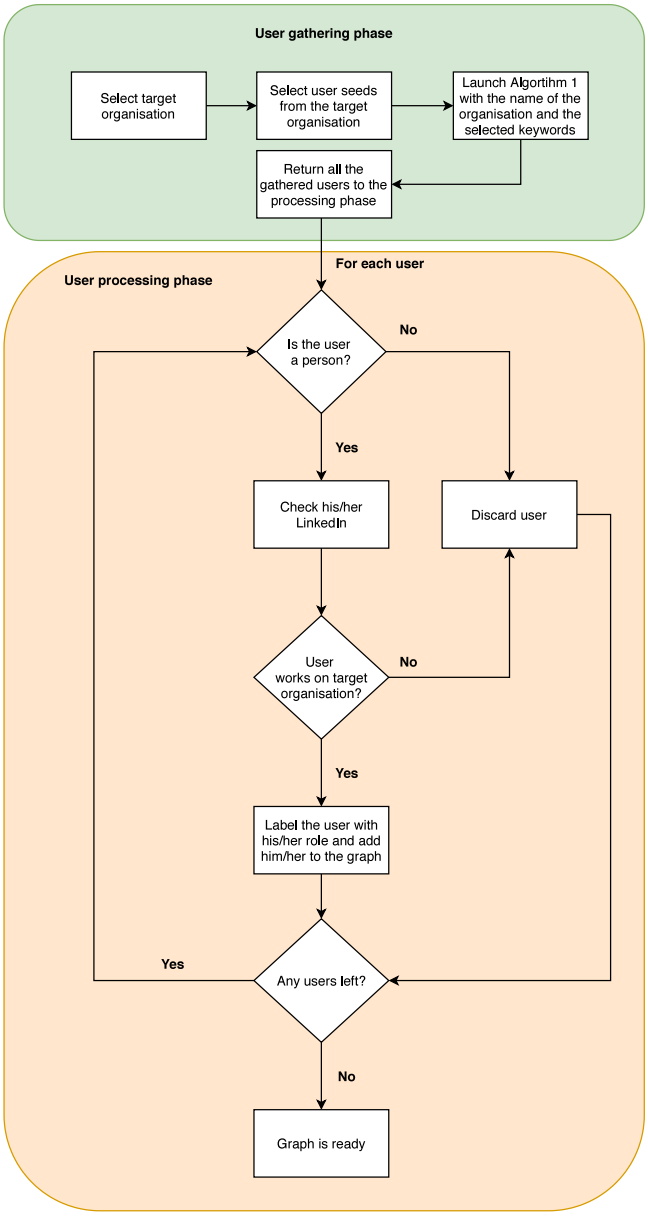


Fig. 1. Methodology for the data collection. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

nodes in the graph. As different centralities model different information, they have been used to model simple behaviour or combined to model more complex information or behaviour within the network itself. Degree-based centralities focus on the links from and to the target node. This centrality has been used to separate fraudsters from legitimate users of an online auction by analysing how fraudsters collude more with each other to increase the price of items (Bangcharoensap et al., 2015). Closeness centrality is calculated by considering the mean distance from the target node to the others in the networks. This centrality identifies nodes that can spread, control or acquire information from the networks; thus, it has been used to map networks of terrorist cells (Krebs, 2002). Betweenness centrality, which also uses the global information of the network, calculates the shortest paths of every pair of nodes and checks on how many paths every node is present. Due to the urge of searching short paths on telecommunications, this centrality has been used to increase the efficiency on this scenario (Borgatti, 2005). Eigenvector and PageRank centralities measure the connectivity of the target node by considering the importance of the nodes with whom it is connected. Eigenvector focuses only on the nodes directly connected with the target node, Pagerank instead conducts random walks over the neighbours of the target node to compute is centrality. Both centralities are used to present recommendations of other accounts to Twitter users by using shared interests and common connections (Gupta et al., 2013).

As we have presented in the previous paragraph, centralities hold information about the nodes, their relationships and the topology of the network; therefore we believe that they fit as features for machine learning classifiers in this task. We will group the centralities into three categories: *Local-based*, *Global-based* and *Katz-based*. Local-based centralities (*Degree, Indegree and Outdegree*) are based on the edges going from and to the target node. Instead, global-based centralities (*Closeness and Betweenness*) include relationships from all the nodes in the graph and the edges that point to and from the target node in their calculations. Finally, Katz-based centralities (*Eigenvector, HITS and PageRank*) are different implementations of the original Katz centrality (Katz, 1953) that is based on the eigenvalues and the degree values.

We now present each of the centralities, for a graph i ($C_*(i)$), that we have used in our approach with their normalised formulas:

- *Degree Centrality* (C_d). This centrality is defined as

$$C_d(i) = \frac{k_i}{N-1}$$

where k_i is the degree (number of edges connected to a node) of node i and N is the total number of nodes in the graph (Opsahl et al., 2010).
- *Indegree Centrality* (C_{In}). This centrality is defined as

$$C_{In}(i) = \frac{k_i}{N-1}$$

where k_i is the number of edges directed to node i, and N is the total number of nodes in the graph (Opsahl et al., 2010).
- *Outdegree Centrality* (C_{Out}). This centrality is defined as

$$C_{Out}(i) = \frac{k_i}{N-1}$$

where k_i is the number of edges directed from node i, and N is the total number of nodes in the graph (Opsahl et al., 2010).

- *Closeness Centrality (C_c)*. This centrality is defined as

$$C_c(i) = \frac{N-1}{\sum_j d(i,j)}$$

where i is the starting node, j the target node and $d(i,j)$ is the distance between them. This measures the distance from the starting node to other nodes in the graph (Newman, 2010).

- *Betweenness Centrality (C_b)*. This centrality is defined as

$$C_b(i) = \sum_{s \neq t \neq i} \frac{\sigma_{st}(i)}{\sigma st} \qquad C_b^*(i) = \frac{2C_b(i)}{(n-1)(n-2)}$$

where i, s and t are three different nodes, σ_{st} are the number of shortest paths from node s and t and $\sigma_{st}(i)$ the number of shortest paths from node s and t that pass through i. This centrality measures the number of shortest paths that pass through a node (Newman, 2010).

- *Eigenvector Centrality (C_e)*. This centrality is defined as

$$C_e(i) = v_i = \frac{1}{\lambda} \sum_j A_{ij} v_j$$

where λ is the eigenvalue and A the adjacency matrix. This is an extension of the degree centrality that rewards the nodes for every connection they have with other individuals in the network (Newman, 2016).

- *HITS Centrality (C_H)*. This centrality consists of two centralities, authority (x_i) and hub (y_i). This centrality are defined as

$$x_i = \alpha \sum_j A_{ij} y_j \qquad y_i = \beta \sum_j A_{ji} x_j$$

where A_{ij} is the connection from the initial node to the rest of nodes that are connected with it and A_{ji} vice versa. Authority measures if a node contains essential information. Hubs, instead, measures if a node points to nodes with high authority in the graph (Kleinberg, 1999).

- *PageRank Centrality (C_p)*. This centrality is defined as

$$C_p(i) = \alpha \sum_j A_{ij} \frac{x_j}{k_j^{out}} + \beta_i$$

where α is the attenuation factor, β a constant, A is the adjacency matrix, k_j^{out} is the outdegree centrality and x_j the value of the centrality for the node j. This centrality represents the connection a node has with other important nodes. The importance assigned to each node is divided among them (Brin and Page, 1998).

4. Evaluation

In this section, we expose the experiments we performed and the results obtained in them. We first describe the setup of the environment where we performed the experiments, and later we explain the experiments and the results obtained.

4.1. Setup

In this study we have conducted several experiments to analyse how the topology of the OSN helps to infer the positions of the employees in the hierarchy of an organisation. We have used the supervised learning algorithms listed below: *Nearest Neighbours* (kNN) (Altman, 1992), *Gaussian Naive Bayes* (NB) (Hand and Yu, 2001), *Decision Tree Classifier* (DTC), *Random Forest* (RF) (Shalev-Shwartz and Ben-David, 2014), *Multi-layer Perceptron* (MLP) (Rumelhart et al., 1985) and *Graph Convolutional Networks* (GCN) (Kipf and Welling, 2016). *kNN, NB, DTC* and *RF* are used to follow with the methodology proposed by Fire and Puzis (2016).

To ensure the reliability of our experiments, we followed the Stratified 10-Folds cross-validation method in each experiment and calculated the average for the results of all the folds. In each of the folds, the train and test sets contain 90% and 10% of the nodes of the dataset respectively, and they follow the same distribution as the whole dataset. Then, we used the set of centralities explained in Section 3 to test the proposed approaches. The centralities were calculated using *NetworkX* (Hagberg et al., 2008) package for Python. All the models, except the GCN, were implemented with the *Scikit-learn* (Pedregosa et al., 2011) package developed for Python. Most of the models use their default parameters except the following ones; in RF, we set the *n_estimators* = 5; in DTC, we set the *min_samples_leaf* = 8 and in kNN, we set the *n_neighbours* to 1, 3, 10 in different experiments. For all the experiments presented in this paper, we calculate the weighted *F-score* metric in the test set of each dataset.

4.2. Experiments & results

In this study, we performed a set of experiments to validate our approach and analyse the performance of each algorithm with the set of proposed centralities. The first experiment compares the performance of our approach (the use of directed graphs and the centralities explained in Section 3.2) with the one proposed by Fire and Puzis. The next experiments were part of an ablation study conducted to the centralities to analyse how they contribute to the classification process. Finally, the last experiments were done to validate how the proposed approach performs in a generalisation problem and a semi-supervised scenario.

In the first experiment, we used multiple supervised machine learning classifiers to measure how they perform with our approach and Fire and Puzis' one. After analysing the results presented in Table 5 we can affirm that the directed graph and the set of centralities proposed in this paper perform better than previously proposed approach to solve this task. We will discuss the results of the experiments in the next section.

To ensure which centralities contribute more to classify nodes, we carried out an ablation study divided into two phases. An ablation study, in this field, consists of eliminating features of a dataset or the algorithm to see which of them are the ones that contribute the most to carry out the proposed task. For this study, we used the algorithm that obtained the best results in the supervised experiment (Random Forest). Afterwards, we tested it with each centrality in the dataset in an isolated way. The results obtained by this study are presented in Table 3.

The second phase of the ablation study involves which class (executives, managers and employees) employees could belong to. To obtain the results of each class, we classified them as a binary classification problem. Through this experiment, we could verify how each of the proposed centralities is contributing when classifying each role. The results of this ablation test are presented in Table 6. In this table, the results show the obtained performance of the classifier per class and centrality in each dataset.

Finally, we conducted the last pair of experiments to test how our approach performs when using smaller quantities of data for the training phase and if it can generalise the information to classify unseen graphs.

For the generalisation experiment, we trained the classifier with graphs from two organisation and tested it with the remaining one. We conducted three experiments and calculated the average result to compare our approach with the results from Fire and Puzis. The results obtained from these test are presented in Table 4. Because of the ease with which this data can be collected from any social network, it is possible to generalise this approach to continue using this algorithm to infer information about the hierarchy of organisations in other social networks.

Moreover, we have carried out a semi-supervised experiment using small amounts of annotated data in order to analyse whether this approach is able to achieve good results when annotated data are not sufficient. To this end, we conducted tests using different percentage of data for the training phase. As shown in Fig. 2, we trained the classifier with 2%, 5% and 10% of the whole data and used the rest for the test part.

Table 3
F1 results of the ablation study conducted to each of the centralities for each of the organisations.

Dataset	Deg.	Indeg.	Outdeg.	Clos.	Betw.	Eigen.	Auth.	Hubs	Page.
Organisation A	65.57	62.24	69.71	65.39	64.07	62.71	66.19	62.74	67.23
Organisation B	51.54	54.50	54.12	56.59	53.92	57.68	52.47	50.15	52.43
Organisation C	75.35	73.90	76.19	74.89	72.70	74.23	77.08	72.94	73.41
AVERAGE	64.16	63.55	**66.67**	65.62	63.56	64.87	65.25	61.94	64.36

Table 4
F1 results for generalisation experiments. The organisation represents the one used as test set.

	Ours	Fire and Puzis
Organisation A	67.47	**67.97**
Organisation B	**57.26**	47.28
Organisation C	**74.26**	62.57
AVERAGE	**66.33**	59.27

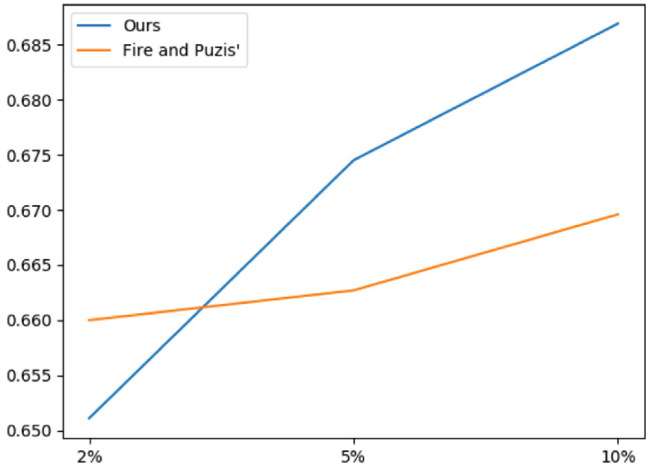


Fig. 2. Results for semi-supervised experiments. In X axis the percentage of the whole dataset used for train is shown and in the Y axis the F1 score is represented.

5. Discussion

The results obtained in the experiments presented in the previous section show that our approach is more effective than the one proposed by Fire and Puzis in classifying the position of an individual in the hierarchy of an organisation present in Twitter. The performance of many supervised machine learning classifiers is presented in Table 5. As the table depicts our proposed set of centralities and the directional graph outperform the work done by Fire and Puzis (2016). Analysing the results shown in Table 5 we can ensure that most of the algorithms tested in the experiment with the directional graphs and the set of centralities in our approach perform better than Fire and Puzis proposal, being Random Forest(RF) the best one.

The main contribution presented in this paper is the analysis conducted on node centralities when classifying the different roles in the organisation. Tables 3 and 6 show the performance obtained in the ablation studies divided per class and dataset separately. We also carried combinatorial experiments testing all the centrality combinations and conclude that none of them makes a significant difference from the others; this is why we did not present the results in the paper.

In the first ablation study, we aimed to analyse which centrality offers more information to classify users in their roles. As Table 3 depicts, the *outdegree* centrality slightly outperforms the results obtained by the rest of them. Delving deeper in the results obtained we can compare the performance of two groups of centralities that express concepts of the same scope: *indegree* vs *outdegree* and *closeness* vs *betweenness*. Analysing the degree centralities we see that the *outdegree* centrality is more important than the *indegree* centrality. We argue that the users that a user follows (followees), represented by *outdegree* centrality, offer more information about their role than the people they are followed by (followers), represented by *indegree* centrality.

On the other hand, *closeness* and *betweenness* measure the distance from the origin node to every other node in the graph and how a node performs acting as a bridge among the others respectively. In this case, we can see how having short paths to every other node on the graph is more important to classify a user than the information of how they act as bridges between nodes. We suspect that this phenomenon occurs because people with strong influence in the organisation and, hence, executives have the greatest *closeness* centrality in the graph.

Table 6 shows the performance of the RF algorithm using centralities individually to classify all the roles separately as a binary problem. Analysing the centralities by categories, we can conclude, by taking the highest performance, that to classify the roles for the employees and managers Katz-based centralities are the best ones. Instead, for the executives, the global centralities are the ones that obtained the best performance and the local-based ones the lowest performance. We believe that this is because executives do not have collective behaviour when following or be followed by other users; therefore we argue that for the inference of information of this user we need to use the global centralities.

Analysing the centralities individually, we can observe that they behave differently for each role. Although we stated before that the global-based centralities are the ones that achieve better performance when classifying executives, the betweenness centrality is one of the worst for this task; instead, is one of the best for classifying the employees. We believe that this may be because the employees are identified with a low betweenness centrality, and the executives' ones vary. Moreover, as executives used to be connected with more people in the organisation, they usually have high closeness centrality, this is why it is the best centrality for this role.

In addition to this, the most notable result is that the *indegree* centrality achieves 0 F-score for every dataset when classifying executives. We assume, as we stated before, that these users do not have a collective behaviour when being followed by users, this is why this centrality may vary between them. Finally, we can also observe that *hubs* centrality does not work on executive users. We consider that this is given because this centrality represents the capability of the node to be connected with authority nodes that, in this case, and depending on the organisation, they can be themselves.

Finally, we conducted two experiments to test how our approach performs in generalisation or semi-supervised scenarios compared with the approach proposed by Fire and Puzis. The results presented in Table 4 show that our approach can be generalised and therefore could be used to classify more organisations without having to be continuously trained, which would allow us to drop the ground truth search from the pipeline. In this way we can extract the data of the users of an organisation from any social network, transform it into a network, extract the centralities of the nodes and classify them with our system.

In the semi-supervised experiment, which results are presented in Fig. 2, our proposed approach performs slightly better than the previously proposed approach except in the first experiment with 2% of the data for the training phase. Thus we consider that the proposed set of centralities and the directed graph work better with larger quantities of data. Moreover, in the other two experiments, our approach rapidly

Table 5
F1 results per algorithm and organisation with our proposed method and the one presented by Fire and Puzis (2016).

	Dataset	KNN (1)	KNN (3)	KNN (10)	NB	DTC	RF	MLP	GCN
Fire	Organisation A	65.78	68.06	68.89	67.87	67.11	69.37	68.28	58.77
	Organisation B	54.52	57.34	56.86	50.70	56.9	59.13	59.03	58.06
	Organisation C	68.52	76.48	77.67	71.86	75.31	77.09	75.25	72.82
	AVERAGE	62.94	**67.29**	67.81	63.48	66.44	68.53	**67.52**	63.22
Ours	Organisation A	67.18	69.65	71.14	69.06	72.45	76.14	69.93	67.70
	Organisation B	52.37	55.63	60.16	52.08	59.12	59.99	55.13	59.51
	Organisation C	73.63	73.37	74.39	71.60	75.57	75.86	72.80	75.37
	AVERAGE	**64.39**	66.22	**68.56**	**64.25**	**69.05**	**70.67**	65.95	**67.53**

Table 6
F1 results per organisation and class for each centrality with *Random Forest* machine learning algorithm.

Dataset	Class	Deg.	Indeg.	Outdeg.	Clos.	Betw.	Eigen.	Auth.	Hubs.	Page.
Organisation A	employee	82.37	81.90	**84.10**	76.99	78.54	75.24	78.71	75.56	79.62
	manager	25.82	14.62	34.66	36.24	30.29	34.56	36.47	30.87	**39.50**
	executive	9.05	0	18.13	21.90	2.50	5.33	17.66	0	**25.53**
Organisation B	employee	63.75	65.41	66.58	67.91	67.70	**70.03**	65.14	63.99	65.49
	manager	44.63	**51.81**	46.61	49.41	44.36	49.48	45.81	40.95	44.35
	executive	14.60	0	6.94	**27.49**	16.19	15.33	13.06	8.78	15.69
Organisation C	employee	85.46	85.60	86.39	84.33	83.84	83.83	**86.58**	84.25	84.51
	manager	49.04	50.63	52.80	55.38	49.26	54.37	**58.42**	49.60	49.67
	executive	16.67	0	6.67	**26.67**	6.67	11.67	2.00	0	13.33

improves its performance due to the information that the directed graph and the centralities provide.

Regarding the applicability of this method, we believe that the experiments carried out and the results obtained demonstrate the ease with which sensitive data can be extracted from people or organisations on Twitter. Although our analysis is not intended to be used to cause any harm to organisations on Twitter, this method can be applied by malicious users to conduct an attack directed at an organisation by looking for members with higher positions in the hierarchy. Furthermore, such malicious users can also extract information from the organisation and build a more detailed profile in order to impersonate one of its members. However, this method can also be used by organisations to detect which of their members are influential in the organisation or to discover new relationships that are unknown to the organisations and thus create new teams within the company itself.

With the experiments we conducted in the previous section, we fulfilled the objective of analysing how the centralities obtained from the nodes in the directed graph contribute to the inference of user roles in organisations. Through the experiments, we have demonstrated that the directed graph with our proposed set of centralities performs better for this task. Furthermore, we also contribute with a theoretical analysis that can be applied in friend-based inference approaches.

6. Conclusions & future work

In this paper, we presented a friend-based inference method that deduces the hierarchical roles of employees in organisations using their relationships on Twitter. To this end, we used data scraped from Twitter to create directed graphs that allow us to use different centralities and extract more information about the relationships of the employees in the OSN. We used these features to classify each employee into one of the three roles proposed, thus increasing the classification roles from the previously proposed approach by Fire and Puzis and using supervised learning methods such as Random Forest.

As discussed in Section 5, we present an in-depth analysis conducted on the proposed set of centralities individually to ascertain how they perform when classifying user into their roles. In this analysis, we realise that the *outdegree* centrality is the one that achieves the better performance and *hubs* centrality the worst. We also did experiments to analyse the performance of every centrality and their categories per class. Through this analysis, we noticed that the *indegree*, and *hubs* centralities perform poorly for the executive users.

Regarding performance in different scenarios, our method has turned to achieve considerably good results when reducing the training set, proving that it can be used with a small quantity of annotated data. Besides, we also tested the performance of our algorithm when testing in an unseen dataset; and this will allow us to use it for new organisations without having to retrain it each time.

Therefore, in our future work, we aim to generalise this analysis to the inference of more attributes so it can be used for any friend-based approaches. Besides, we will increase the datasets by collecting more users or using the algorithm presented in this paper to capture new organisation data. Also, we will improve the analysis by mixing behavioural data from the users to infer new attributes and relationships. Moreover, the novel graph centred methods, such as GCN, performed excellent solving other tasks but not that well in this one. Thus, we will consider in working to improve and adapt these models to outperform the classic algorithms in this task and the ones related to it.

CRediT authorship contribution statement

Rubén Sánchez-Corcuera: Methodology, Software, Validation, Writing - original draft. **Aritz Bilbao-Jayo:** Resources, Data curation, Writing - review & editing. **Unai Zulaika:** Software, Validation, Writing - review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We gratefully acknowledge the support of the Basque Governments Department of Education, Spain for the predoctoral funding of some of the authors and for DEUSTEK4: Entornos inteligentes abiertos y tecnologías para el aprendizaje, recognised group of the Basque University System, IT1078-16. We also gratefully acknowledge the support of NVIDIA Corporation, Spain for the donation of the hardware used in this research. Finally, we also wanted to acknowledge the Weight and Biases (Biewald, 2020) team for their service and support.

References

Altenburger, K.M., Ugander, J., 2018. Monophily in social networks introduces similarity among friends-of-friends. Nat. Human Behav. 2, 284–290.

Altman, N.S., 1992. An introduction to kernel and nearest-neighbor nonparametric regression. Amer. Statist. 46, 175–185. http://dx.doi.org/10.2307/2685209.

Bangcharoensap, P., Kobayashi, H., Shimizu, N., Yamauchi, S., Murata, T., 2015. Two step graph-based semi-supervised learning for online auction fraud detection. In: Joint European Conference on Machine Learning and Knowledge Discovery in Databases. Springer, pp. 165–179.

Bastos, M.T., Mercea, D., 2019. The brexit botnet and user-generated hyperpartisan news. Soc. Sci. Comput. Rev. 37, 38–54. http://dx.doi.org/10.2139/ssrn.3034051.

Biewald, L., 2020. Experiment tracking with weights and biases. URL: https://www.wandb.com/, software available from wandb.com.

Bilbao-Jayo, A., Almeida, A., 2018. Political discourse classification in social networks using context sensitive convolutional neural networks. In: Proceedings of the Sixth International Workshop on Natural Language Processing for Social Media. pp. 76–85. http://dx.doi.org/10.18653/v1/w18-3513.

Borgatti, S.P., 2005. Centrality and network flow. Soc. Netw. 27, 55–71.

Bovet, A., Makse, H.A., 2016. Influence of fake news in twitter during the 2016 us presidential election. Nat. Commun. 10, 7. http://dx.doi.org/10.1038/s41467-018-07761-2.

Brin, S., Page, L., 1998. The anatomy of a large-scale hypertextual web search engine. Comput. Netw. ISDN Syst. 30, 107–117.

Chaabane, A., Acs, G., Kaafar, M.A., et al., 2012. You are what you like! information leakage through users' interests. In: Proceedings of the 19th Annual Network & Distributed System Security Symposium. NDSS, Citeseer.

Chaniotakis, E., Antoniou, C., Aifadopoulou, G., Dimitriou, L., 2017. Inferring activities from social media data. Transp. Res. Rec. 2666, 29–37.

Chen, J., He, J., Cai, L., Pan, J., 2016. Profiling online social network users via relationships and network characteristics. In: 2016 IEEE Global Communications Conference. GLOBECOM, IEEE, pp. 1–6. http://dx.doi.org/10.1109/glocom.2016.7842176.

Davis, C.A., Varol, O., Ferrara, E., Flammini, A., Menczer, F., 2016. Botornot: A system to evaluate social bots. In: Proceedings of the 25th International Conference Companion on World Wide Web. International World Wide Web Conferences Steering Committee, pp. 273–274. http://dx.doi.org/10.1145/2872518.2889302.

Diallo, S.Y., Lynch, C.J., Gore, R., Padilla, J.J., 2016. Identifying key papers within a journal via network centrality measures. Scientometrics 107, 1005–1020.

Fani, H., Jiang, E., Bagheri, E., Al-Obeidat, F., Du, W., Kargar, M., 2019. User community detection via embedding of social network structure and temporal content. Inf. Process. Manage. 102056. http://dx.doi.org/10.1016/j.ipm.2019.102056.

Ferrara, E., Varol, O., Davis, C., Menczer, F., Flammini, A., 2016. The rise of social bots. Commun. ACM 59, 96–104. http://dx.doi.org/10.1145/2818717.

Fire, M., Puzis, R., 2016. Organization mining using online social networks. Netw. Spat. Econ. 16, 545–578. http://dx.doi.org/10.1007/s11067-015-9288-4.

Garton, L., Haythornthwaite, C., Wellman, B., 1997. Studying online social networks. J. Comput.-Mediat. Commun. 3, JCMC313.

Gong, N.Z., Liu, B., 2016. You are who you know and how you behave: Attribute inference attacks via users' social friends and behaviors. In: 25th USENIX Security Symposium, USENIX Security 16. pp. 979–995.

Gong, N.Z., Liu, B., 2018. Attribute inference attacks in online social networks. ACM Trans. Priv. Secur. 21, 1–30. http://dx.doi.org/10.1145/3154793.

Grinberg, N., Joseph, K., Friedland, L., Swire-Thompson, B., Lazer, D., 2016. Fake news on twitter during the 2016 us presidential election. Science 363, 374–378. http://dx.doi.org/10.1126/science.aau2706.

Grunspan, D.Z., Wiggins, B.L., Goodreau, S.M., 2014. Understanding classrooms through social network analysis: A primer for social network analysis in education research. CBE Life Sci. Educ. 13, 167–178.

Guimera, R., Mossa, S., Turtschi, A., Amaral, L.N., 2005. The worldwide air transportation network: Anomalous centrality, community structure, and cities' global roles. Proc. Natl. Acad. Sci. 102, 7794–7799.

Gupta, P., Goel, A., Lin, J., Sharma, A., Wang, D., Zadeh, R., 2013. Wtf: The who to follow service at twitter. In: Proceedings of the 22nd International Conference on World Wide Web. pp. 505–514.

Hagberg, A., Swart, P., Chult, D.S., 2008. Exploring Network Structure, Dynamics, and Function using NetworkX. Technical Report, Los Alamos National Lab (LANL), Los Alamos, NM (United States).

Hand, D.J., Yu, K., 2001. Idiot's bayes—not so stupid after all? Internat. Statist. Rev. 69, 385–398. http://dx.doi.org/10.1111/j.1751-5823.2001.tb00465.x.

Howard, B., 2008. Analyzing online social networks. Commun. ACM 51, 14–16.

Jayaweera, I., Perera, K., Munasinghe, J., 2017. Centrality measures to identify traffic congestion on road networks: A case study of Sri Lanka. IOSR J. Math. 13, 13–19.

Jia, J., Wang, B., Zhang, L., Gong, N.Z., 2017. Attriinfer: Inferring user attributes in online social networks using markov random fields. In: Proceedings of the 26th International Conference on World Wide Web. pp. 1561–1569.

Katz, L., 1953. A new status index derived from sociometric analysis. Psychometrika 18, 39–43. http://dx.doi.org/10.1007/bf02289026.

Kim, J., Hastak, M., 2018. Social network analysis: Characteristics of online social networks after a disaster. Int. J. Inf. Manage. 38, 86–96.

Kipf, T.N., Welling, M., 2016. Semi-supervised classification with graph convolutional networks. arXiv preprint arXiv:1609.02907.

Kleinberg, J.M., 1999. Hubs, authorities, and communities. ACM Comput. Surv. 31, 5-es.

Koschützki, D., Lehmann, K.A., Peeters, L., Richter, S., Tenfelde-Podehl, D., Zlotowski, O., 2005. Centrality indices. In: Network Analysis. Springer, pp. 16–61.

Krebs, V.E., 2002. Mapping networks of terrorist cells. Connections 24, 43–52.

Li, R., Wang, S., Deng, H., Wang, R., Chang, K.C.-C., 2012. Towards social user profiling: Unified and discriminative influence model for inferring home locations. In: 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD 2012. pp. 1023–1031.

McPherson, M., Smith-Lovin, L., Cook, J.M., 2001. Birds of a feather: Homophily in social networks. Annu. Rev. Sociol. 27, 415–444. http://dx.doi.org/10.1146/annurev.soc.27.1.415.

Mei, B., Xiao, Y., Li, H., Cheng, X., Sun, Y., 2017. Inference attacks based on neural networks in social networks. In: Proceedings of the Fifth ACM/IEEE Workshop on Hot Topics in Web Systems and Technologies. ACM, p. 10.

Mulders, D., De Bodt, C., Bjelland, J., Pentland, A., Verleysen, M., de Montjoye, Y.-A., 2019. Inference of node attributes from social network assortativity. Neural Comput. Appl. 1–21.

Newman, M., 2010. Networks: An Introduction. Oxford University Press.

Newman, M.E., 2016. Mathematics of Networks. Springer.

Noulas, A., Mascolo, C., Frias-Martinez, E., 2013. Exploiting foursquare and cellular data to infer user activity in urban environments. In: 2013 IEEE 14th International Conference on Mobile Data Management, Vol. 1. IEEE, pp. 167–176.

Opsahl, T., Agneessens, F., Skvoretz, J., 2010. Node centrality in weighted networks: Generalizing degree and shortest paths. Soc. Netw. 32, 245–251. http://dx.doi.org/10.1016/j.socnet.2010.03.006.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, E., 2011. Scikit-learn: Machine learning in Python. J. Mach. Learn. Res. 12, 2825–2830.

Rumelhart, D.E., Hinton, G.E., Williams, R.J., 1985. Learning Internal Representations by Error Propagation. Technical Report, California Univ San Diego La Jolla Inst for Cognitive Science, http://dx.doi.org/10.21236/ada164453.

Saltz, J.S., Hiltz, S.R., Turoff, M., 2004. Student social graphs: visualizing a student's online social network. In: Proceedings of the 2004 ACM Conference on Computer Supported Cooperative Work. pp. 596–599.

Shalev-Shwartz, S., Ben-David, S., 2014. Decision trees. In: Understanding Machine Learning. Cambridge University Press, pp. 250–256. http://dx.doi.org/10.1017/cbo9781107298019.004.

Steinfield, C., Ellison, N.B., Lampe, C., 2008. Social capital, self-esteem, and use of online social network sites: A longitudinal analysis. J. Appl. Dev. Psychol. 29, 434–445.

Thomas, K., Grier, C., Nicol, D.M., 2010. unfriendly: Multi-party privacy risks in social networks. In: International Symposium on Privacy Enhancing Technologies Symposium. Springer, pp. 236–252.

Valverde-Rebaza, J.C., Roche, M., Poncelet, P., de Andrade Lopes, A., 2018. The role of location and social strength for friendship prediction in location-based social networks. Inf. Process. Manage. 54, 475–489. http://dx.doi.org/10.1016/j.ipm.2018.02.004.

Varol, O., Ferrara, E., Davis, C.A., Menczer, F., Flammini, A., 2017. Online human-bot interactions: Detection, estimation, and characterization. In: Eleventh International AAAI Conference on Web and Social Media. pp. 280–289.

You, Q., Bhatia, S., Sun, T., Luo, J., 2014. The eyes of the beholder: Gender prediction using images posted in online social networks. In: 2014 IEEE International Conference on Data Mining Workshop. IEEE, pp. 1026–1030.

Zarrinkalam, F., Kahani, M., Bagheri, E., 2018. Mining user interests over active topics on social networks. Inf. Process. Manage. 54, 339–357. http://dx.doi.org/10.1016/j.ipm.2017.12.003.

Zhan, J., Gurung, S., Parsa, S.P.K., 2017. Identification of top-k nodes in large networks using katz centrality. J. Big Data 4, 1–19.

Zhang, J., Hu, X., Zhang, Y., Liu, H., 2016. Your age is no secret: Inferring microbloggers' ages via content and interaction analysis. In: Tenth International AAAI Conference on Web and Social Media. IEEE, pp. 476–485.