Detection of Terrorism-related Twitter Communities using Centrality Scores

Ilias Gialampoukidis, George Kalpakis, Theodora Tsikrika, Symeon Papadopoulos, Stefanos Vrochidis and Ioannis Kompatsiaris
Information Technologies Institute
Centre for Research and Technology Hellas
Thessaloniki, Greece 57001
{heliasgj,kalpakis,theodora.tsikrika,papadop,stefanos,ikom}@iti.gr

ABSTRACT

Social media are widely used among terrorists to communicate and disseminate their activities. User-to-user interaction (e.g. mentions, follows) leads to the formation of complex networks, with topology that reveals key-players and key-communities in the terrorism domain. Both the administrators of social media platforms and Law Enforcement Agencies seek to identify not only single users but groups of terrorism-related users so that they can reduce the impact of their information exchange efforts. To this end, we propose a novel framework that combines community detection with key-player identification to retrieve communities of terrorism-related social media users. Experiments show that most of the members of each retrieved key-community are already suspended by Twitter, violating its terms, and are hence associated with terrorism-oriented content with high probability.

CCS CONCEPTS

• **Information systems** → **Information retrieval**; **Test collections**; *Web searching and information discovery*; *Multimedia and multimodal retrieval*; • **Human-centered computing** → **Social networking sites**; • **Networks** → *Social media networks*;

KEYWORDS

Social Network Analysis, key-player identification, community detection, terrorism-oriented social media mining, Twitter

ACM Reference format:
Ilias Gialampoukidis, George Kalpakis, Theodora Tsikrika, Symeon Papadopoulos, Stefanos Vrochidis and Ioannis Kompatsiaris. 2017. Detection of Terrorism-related Twitter Communities using Centrality Scores. In *Proceedings of MFSec'17, Bucharest, Romania, June 06, 2017,* 5 pages.
https://doi.org/10.1145/3078897.3080534

MFSec'17, June 06, 2017, Bucharest, Romania

ACM ISBN 978-1-4503-5034-1/17/06…$15.00
https://doi.org/10.1145/3078897.3080534

1 INTRODUCTION

The rapid growth of the Internet has resulted in modern forms of communication and exchange of information, realized mainly through the use of social media networking platforms (e.g. Twitter, Facebook, etc.), which have dominated the online world during the past few years. Social media networks have made possible the communication among people across nationalities, religions, cultures or residences; however, their great power and reach has become an attractive feature for their use by terrorist and extremist organizations for disseminating their propaganda, recruiting and radicalizing new members, raising funds, organizing operations, and publishing information and instructions exploited by lone-wolf terrorists when preparing and committing acts of terror [27–29].

Due to its nature that permits the inexpensive communication of multimedia messages (i.e. tweets) to users worldwide, Twitter has been used primarily for promoting and spreading their propaganda typically using a top-down approach, with a core group of members spreading the group's messages, which are then re-shared by other affiliated accounts. Both the administrators of the social media networking platform itself (Twitter), on the one hand, and the Law Enforcement Agencies (LEAs), on the other, are interested in monitoring terrorism-related activities taking place through the platform. In the former case, the goal is to detect material that violates the platform's terms and conditions regarding extremist content, while in the latter case such information may be very useful in investigations for prosecuting the perpetrators of terrorist attacks. In both cases, it is of vital significance to detect the communities in the social networks and their most prominent users (i.e. key players) who disseminate terrorism-related information, so as to prevent terrorist groups from spreading their propaganda (to the extent possible), by shutting down accounts who are found to play a central role in this information exchange.

Over the past two decades, several research efforts have discussed the network structure of terrorist organizations. One of the early efforts examined the network structure of the 9/11 hijackers along with their accomplices and detected the ring leaders of the terrorist attacks based on their social associations [15]. Later work focused on using social network analysis for examining the basic characteristics of terrorist groups or organizations [26]. More recent research has examined the survival mechanisms of the Global Salafi Jihad (GSJ) terrorist network, even after being severely damaged by the authorities, by analyzing its network structure and topology [30]. In addition, several works have been conducted for studying the use of social media, and especially Twitter, by terrorist

organizations. Specifically, a work has examined the significant role of Twitter in facilitating terrorists to execute their attack in Mumbai (November 2008), by monitoring and exploiting situational information which was broadcast through Twitter [19]. More recent research has studied the Islamic State's (IS) strategy for communicating their propaganda for radicalizing and recruiting Twitter users [6]. Furthermore, the significant role played by feeder accounts of terrorist organizations for exchanging information from the Syria insurgency zone is pointed out in [14]. Key player identification in complex networks, on the other hand, has been mainly addressed through the use of different centrality measures; e.g. recent work [10] has used several centrality measures to rank terrorism-related Twitter accounts based on their location in the network and the topology of the network of user-to-user mentions.

This work aims at identifying groups of terrorism-related users exchanging information through social media platforms by detecting the key players of a social media network and the interrelated communities of users interacting with them. To this end, we extend the approach of [10] and propose a hybrid framework which first retrieves the key network players and then enriches the retrieved results by adding the members of a user's detected community based on the combination of centrality scores with community detection algorithms. These centrality measures, which aim to identify key-players in the terrorism domain, are estimated on social media networks based on user mentions and are compared with other popularity measures (i.e. number of followers, number of friends) used for identifying very important users within the structure of these networks. This work also presents a case study on a social media network formed by Twitter accounts based on a set of terrorism-related Arabic keywords provided by LEAs and domain experts, for demonstrating the performance of our proposed framework based on evidence related to the suspension of the majority of the retrieved Twitter accounts.

2 KEY TERRORISM COMMUNITY DETECTION FRAMEWORK

In this work, entropy-based centrality measures are exploited to first retrieve a list of key-players and then a community detection algorithm to enrich the initial set of results. Our framework is presented in Figure 1, where keyword-based search provides a set of social media posts. Based on this, a network of mentions is created, using the user-to-user interactions contained in the corresponding posts. In the resulting network of users, each user is represented by a node and a link between two users (i, k) exists if user n_i mentions or is mentioned by user n_k. We use entropy-based centralities to, first, identify key-players [10] and we then extend the method by associating key-players with their community.

2.1 Centrality-based key player identification

We denote by $G(N, L)$ the network of mentions with N nodes (users accounts) and L links. The network is unweighted and undirected capturing only the user-to-user interactions in Twitter or any other social media domain. The degree of a node n_k is denoted by $deg(n_k)$, and is equal to the number of its adjacent links. The degree is normalized to define the degree centrality as follows [9]:

$$DC_k = \frac{deg(n_k)}{N-1} \tag{1}$$

The degree simply counts the number of nodes and is not affected by the position of a hub in the network. However, the betweenness centrality [9] of a node n_k is based on the number of paths $g_{ij}(n_k)$ from node n_i to node n_j that pass through node n_k, divided by the number of all paths g_{ij} from node n_i to node n_j, summed over all pairs of nodes (n_i, n_j) and normalized by its maximum value:

$$BC_k = \frac{2\sum_{i<j}^{N} \frac{g_{ij}(n_k)}{g_{ij}}}{N^2 - 3N + 2} \tag{2}$$

Nodes with high betweenness centrality are very important for the communication in a network [1] , due to the fact that their removal strongly affects the network connectivity and robustness. Other centrality measures have also been proposed, based on the mutual distances of all nodes (closeness centrality) [9], on the influence of a node (eigenvector centrality) [4], or motivated by the importance of a Web page (PageRank) [5].

In the context of this work, we propose the use of entropy-based centrality measures, such as the Mapping Entropy (ME) and the Mapping Entropy Betweenness (MEB), taking also into account the neighborhood $\mathcal{N}(n_k)$ of a node n_k. *Mapping Entropy* centrality [18] is defined as a function of the degree centrality:

$$ME_k = -DC_k \sum_{n_i \in \mathcal{N}(n_k)} \log DC_i \tag{3}$$

whereas *Mapping Entropy Betweenness* centrality [10] is defined as a function of betweenness centrality:

$$MEB_k = -BC_k \sum_{n_i \in \mathcal{N}(n_k)} \log BC_i \tag{4}$$

Intuitively, to interpret Equations (3) and (4), one may think of a random walker on the network, standing at node n_k, who picks his/her next step with probability DC_i (BC_i). Then, the weight $-\log DC_i$ ($-\log BC_i$) is interpreted as the Shannon information of the event that the random walker picked node n_i, and is summed over all neighbors of node n_k. These two measures consider the information that is communicated through nodes who act as a hub (bridge), i.e. those with high values of degree (betweenness) centrality between any two members. In particular, the MEB centrality considers the betweenness centrality of a node and exploits local information from its neighborhood; hence, high MEB values indicate that a particular node can act as a bridge for disseminating information, even if their degree centrality is low [22].

In the following, we combine the key-player identification methods with community detection approaches that are able to cluster the network into communities of densely connected user accounts.

2.2 Community detection around key players

In parallel to the key-player identification, a community detection algorithm is used to divide the network into groups of users (communities). The top-ranked key-player is used to enrich the retrieved results, which is achieved by searching for the community where the key-player belongs to.

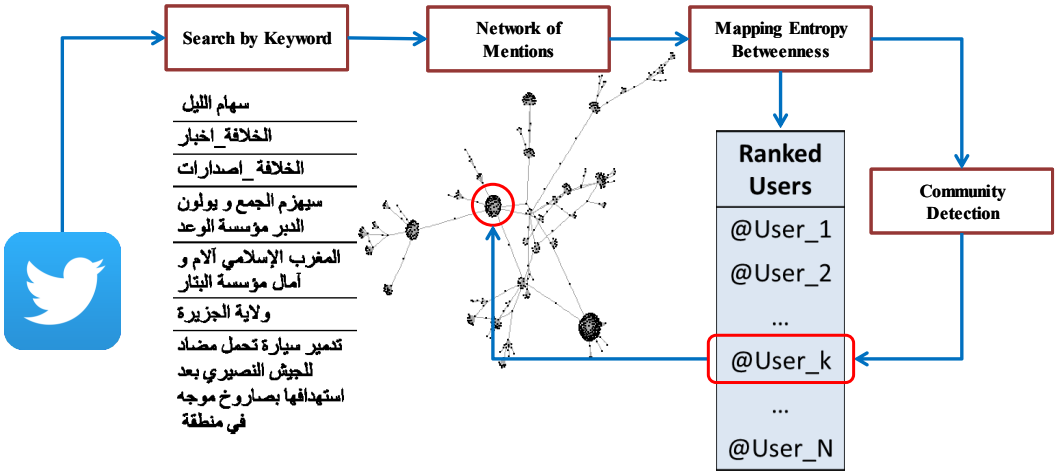

Figure 1: Key terrorism-related community detection on the network of Twitter mentions.

Community detection in complex networks aims to identify groups of nodes that are more densely connected to each other within a group than to the rest of the network outside of the group [20]. The groups are communities of users in the social media domain, sharing a common property or playing similar roles within the network [8]. Community structure is very popular in many fields, including sociology and biology [12], as well as computer science [17], and in any domain where systems or items admit a network representation. Detecting communities in complex networks is often viewed as a graph partitioning problem, where all nodes are assigned to a community, but density-based approaches leave out noise, i.e. do not assign all nodes to communities. In our experiments, we shall present and compare both approaches.

Several community detection algorithms have been proposed (e.g. [2, 8, 12, 13, 16, 21, 23, 25]). The network is partitioned into communities using either the maximization of modularity [2, 17], the minimization of codelength [24] or density-based approaches [11]. We present in the experiments the key-community, defined as the community that the key-player belongs to, as provided by the algorithms FastGreedy [7], Walktrap [21], Infomap [3, 24, 25], Louvain [2] and DBSCAN*-Martingale [11]. The most popular methods are those aiming at the maximization of modularity, defined as [7]:

$$Q = \frac{1}{2m} \sum_{i=1}^{c} (e_{ii} - \alpha_i^2) \tag{5}$$

where e_{ij} is the fraction of links between a node in community i and a node in community j, α_i is the fraction of links between two members of the community i, $m = \sum_k deg(n_k)$, and c is the number of communities. We adopt the modularity maximization community detection approach as a fast and scalable approach that admits hierarchical and iterative methods [2, 20] to maximize the objective function of Equation 5. Assuming the key-player is a member of the k-th community, our framework returns all its members $n_{k_1}, n_{k_2}, \ldots, n_{k_l}$, all of which are marked as the final list of accounts with suspicious activity.

3 EXPERIMENTS

We evaluate our framework in a network consisting of terrorism-related Twitter accounts formed based on user mentions.

As ground-truth we make use of information from Twitter, which marks user accounts as suspended, given that the suspension process is applied when an account violates Twitter rules by exhibiting abusive behavior, including posting content related to violent threats and hate speech (Twitter has suspended 360,000 terrorism-related accounts from mid-2015 until August 2016[1]). Our data were collected by executing queries on the Twitter API[2] based on a set of five Arabic keywords related to terrorist propaganda. These keywords were provided by LEAs and domain experts and are related to the Caliphate State, its news, publications, and photos from the Caliphate area. The collected dataset consists of 9,528 Twitter posts by 4,400 users. The top-100 user accounts are retrieved in the key-player identification step using the ranking methods of Table 1 and are then combined with the community detection approaches of Table 2. The evaluation is performed by assessing whether these accounts are suspended, active or no longer exist (i.e. accounts which have been temporarily or permanently deactivated).

The first part of our framework evaluates several centrality measures, including the proposed Mapping Entropy and Mapping Entropy Betweenness, as well as popularity measures, such as the number of friends and followers, in terms of their ability to retrieve suspended users. The results in Table 1 indicate that the entropy-based centralities ME and MEB are able to retrieve the first suspended user at position 16, while PageRank follows at position 19. Other centrality and popularity measures, such as closeness, eigenvector and number of followers do not find any suspended

[1] https://blog.twitter.com/2016/an-update-on-our-efforts-to-combat-violent-extremism
[2] https://dev.twitter.com/

Table 1: Comparison among several ranking methods

Ranking Method	Position	Reciprocal Rank
Degree centrality	20	5.00%
Betweenness centrality	35	2.86%
Closeness centrality	> 100	< 1.00%
Eigenvector centrality	> 100	< 1.00%
Num of followers	> 100	< 1.00%
Num of friends	31	3.23%
PageRank	19	5.26%
Mapping Entropy	16	6.25%
Mapping Entropy Betweenness	16	6.25%

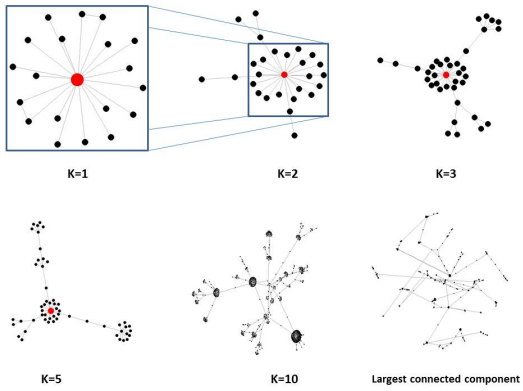

Figure 2: First, second, third, fifth and tenth order neighborhoods of the suspended user and the largest component.

user at the top-100 positions of their retrieved users. We observe that the network is very spread with many bridges and a diameter equal to 27, so key-players are expected to be positioned in between many pairs of nodes in the network, exploiting also their neighborhood's high betweenness centrality.

The K-th order neighborhood $\mathcal{N}_K$ of node n_j is the set of all nodes that are reachable from n_j within $K-1$ intermediate nodes: $\mathcal{N}_K(n_j) = \{n \in N : d(n, n_j) \leq K\}$, where $d(n, n_j)$ is the network distance of any two nodes. In Figure 2 we show the first ($K = 1$), second ($K = 2$), third ($K = 3$), fifth ($K = 5$) and tenth ($K = 10$) order neighborhoods of the first suspended user and the largest connected component. Although the ME and MEB centralities both retrieve a suspended user at rank 16, the user does not correspond to the same Twitter account. In fact, the Twitter user at the 16^{th} position of ME centrality leads to a disconnected component of two users, where one of them is suspended and the other is not. However, the neighborhood of the suspended user (Figure 2) from the MEB centrality is part of the largest connected component of the network with 1,334 accounts. Therefore, we proceed to the next step by considering the MEB centrality measure and not ME.

Given the first identified suspended user in the MEB ranking, we explore the community where the user belongs to. The results are reported in Table 2, along with the community size per community

Table 2: Comparison of community detection methods

Method	Community size	Active users	Do not exist	Suspended accounts
FastGreedy	58	4 (6.9%)	6 (10.3%)	48 (82.8%)
Walktrap	33	3 (9.1%)	4 (12.1%)	26 (78.8%)
Louvain	58	4 (6.9%)	6 (10.3%)	48 (82.8%)
Infomap	20	4 (20.0%)	3 (15.0%)	13 (65.0%)
DBSCAN*-Martingale	23	2 (8.7%)	3 (13.0%)	18 (78.3%)

Figure 3: A sample set of images uploaded by key-players with militaristic or nationalistic content. Faces are redacted so as to avoid the inclusion of sensitive information.

detection method. We observe that in all cases examined, the majority of accounts are already suspended and some of them no longer exist. In particular, the modularity maximization methods (FastGreedy, Louvain) are able to retrieve the largest communities and thus more accounts with potentially illegal activity. The percentage of suspended users is 82.76% for the modularity maximization approaches and 78% for the Walktrap and DBSCAN*-Martingale, indicating a marginal advantage for the former. The community provided by Infomap is very small, compared to the other community sizes, but still the number of active accounts (not yet suspended) is only 20%. Figure 3 depicts sample content from such active accounts that have not been marked as suspended by Twitter. One may note that their content is military-themed, indicating potentially suspicious user activity even in non-suspended accounts.

4 CONCLUSIONS

We proposed a hybrid model that combines MEB centrality and community detection that retrieves groups of social media user accounts that are key-players in the terrorism domain. We found that centrality measures on the network of mentions perform better than other popularity measures (number of followers or friends) in finding key-players in the terrorism domain. Given a terrorism-related user, his/her network community reveals a group of additional terrorism-related users, exploiting the outcome of a community detection method, with modularity maximization methods outperforming density-based and other methods.

ACKNOWLEDGMENTS

This work was supported by the project TENSOR (H2020-700024), funded by the European Commission.

REFERENCES

[1] Ala Berzinji, Lisa Kaati, and Ahmed Rezine. 2012. Detecting key players in terrorist networks. In *Intelligence and Security Informatics Conference (EISIC), 2012 European*. IEEE, 297–302.

[2] Vincent D Blondel, Jean-Loup Guillaume, Renaud Lambiotte, and Etienne Lefebvre. 2008. Fast unfolding of communities in large networks. *Journal of statistical mechanics: theory and experiment* 2008, 10 (2008), P10008.

[3] Ludvig Bohlin, Daniel Edler, Andrea Lancichinetti, and Martin Rosvall. 2014. Community detection and visualization of networks with the map equation framework. In *Measuring Scholarly Impact*. Springer, 3–34.

[4] Phillip Bonacich and Paulette Lloyd. 2001. Eigenvector-like measures of centrality for asymmetric relations. *Social networks* 23, 3 (2001), 191.

[5] Sergey Brin and Lawrence Page. 2012. Reprint of: The anatomy of a large-scale hypertextual web search engine. *Computer networks* 56, 18 (2012), 3825–3833.

[6] Akemi Takeoka Chatfield, Christopher G Reddick, and Uuf Brajawidagda. 2015. Tweeting propaganda, radicalization and recruitment: Islamic state supporters multi-sided twitter networks. In *Proceedings of the 16th Annual International Conference on Digital Government Research*. ACM, 239–249.

[7] Aaron Clauset, Mark EJ Newman, and Cristopher Moore. 2004. Finding community structure in very large networks. *Physical review E* 70, 6 (2004), 066111.

[8] Santo Fortunato. 2010. Community detection in graphs. *Physics reports* 486, 3 (2010), 75–174.

[9] Linton C Freeman. 1978. Centrality in social networks conceptual clarification. *Social networks* 1, 3 (1978), 215–239.

[10] Ilias Gialampoukidis, George Kalpakis, Theodora Tsikrika, Stefanos Vrochidis, and Ioannis Kompatsiaris. 2016. Key player identification in terrorism-related social media networks using centrality measures. In *European Intelligence and Security Informatics Conference (EISIC 2016), August*. 17–19.

[11] Ilias Gialampoukidis, Theodora Tsikrika, Stefanos Vrochidis, and Ioannis Kompatsiaris. 2016. Community detection in complex networks based on DBSCAN* and a Martingale process. In *Semantic and Social Media Adaptation and Personalization (SMAP), 2016 11th International Workshop on*. IEEE, 1–6.

[12] Michelle Girvan and Mark EJ Newman. 2002. Community structure in social and biological networks. *Proceedings of the national academy of sciences* 99, 12 (2002), 7821–7826.

[13] Steve Harenberg, Gonzalo Bello, L Gjeltema, Stephen Ranshous, Jitendra Harlalka, Ramona Seay, Kanchana Padmanabhan, and Nagiza Samatova. 2014. Community detection in large-scale networks: a survey and empirical evaluation. *Wiley Interdisciplinary Reviews: Computational Statistics* 6, 6 (2014), 426–439.

[14] Jytte Klausen. 2015. Tweeting the Jihad: Social media networks of Western foreign fighters in Syria and Iraq. *Studies in Conflict & Terrorism* 38, 1 (2015), 1–22.

[15] Valdis Krebs. 2002. Uncloaking terrorist networks. *First Monday* 7, 4 (2002).

[16] Fragkiskos D Malliaros and Michalis Vazirgiannis. 2013. Clustering and community detection in directed networks: A survey. *Physics Reports* 533, 4 (2013), 95–142.

[17] ME Newman and M Girvan. 2004. Finding and evaluating community structure in networks. *Physical Review E* 69, 2 (2004), 026113.

[18] Tingyuan Nie, Zheng Guo, Kun Zhao, and Zhe-Ming Lu. 2016. Using mapping entropy to identify node centrality in complex networks. *Physica A: Statistical Mechanics and its Applications* 453 (2016), 290–297.

[19] Onook Oh, Manish Agrawal, and H Raghav Rao. 2011. Information control and terrorism: Tracking the Mumbai terrorist attack through twitter. *Information Systems Frontiers* 13, 1 (2011), 33–43.

[20] Symeon Papadopoulos, Yiannis Kompatsiaris, Athena Vakali, and Ploutarchos Spyridonos. 2012. Community detection in social media. *Data Mining and Knowledge Discovery* 24, 3 (2012), 515–554.

[21] Pascal Pons and Matthieu Latapy. 2006. Computing communities in large networks using random walks. *J. Graph Algorithms Appl.* 10, 2 (2006), 191–218.

[22] Jialun Qin, Jennifer J Xu, Daning Hu, Marc Sageman, and Hsinchun Chen. 2005. Analyzing terrorist networks: A case study of the global salafi jihad network. In *Intelligence and security informatics*. Springer, 287–304.

[23] Usha Nandini Raghavan, Réka Albert, and Soundar Kumara. 2007. Near linear time algorithm to detect community structures in large-scale networks. *Physical Review E* 76, 3 (2007), 036106.

[24] Martin Rosvall, Daniel Axelsson, and Carl T Bergstrom. 2009. The map equation. *The European Physical Journal Special Topics* 178, 1 (2009), 13–23.

[25] Martin Rosvall and Carl T Bergstrom. 2008. Maps of random walks on complex networks reveal community structure. *Proceedings of the National Academy of Sciences* 105, 4 (2008), 1118–1123.

[26] Sudhir Saxena, K Santhanam, and Aparna Basu. 2004. Application of social network analysis (SNA) to terrorist networks in Jammu & Kashmir. *Strategic Analysis* 28, 1 (2004), 84–101.

[27] Robin L Thompson. 2011. Radicalization and the use of social media. *Journal of strategic security* 4, 4 (2011), 167.

[28] Robyn Torok. 2010. "Make A Bomb In Your Mums Kitchen": Cyber Recruiting And Socialisation of 'White Moors' and Home Grown Jihadists. (2010).

[29] Ines Von Behr. 2013. Radicalisation in the digital era: The use of the Internet in 15 cases of terrorism and extremism. (2013).

[30] Jie Xu, Daning Hu, and Hsinchun Chen. 2009. The dynamics of terrorist networks: Understanding the survival mechanisms of Global Salafi Jihad. *Journal of Homeland Security and Emergency Management* 6, 1 (2009), 1–15.