SoNA: A Knowledge-based Social Network Analysis Framework for Predictive Policing

Michael Spranger*, Hanna Siewerts*, Joshua Hampl*, Florian Heinke* and Dirk Labudde*†
*University of Applied Sciences Mittweida
Faculty Applied Computer Sciences & Biosciences
Mittweida, Germany
Email: {*name.surname*}@hs-mittweida.de
†Fraunhofer
Cyber Security
Darmstadt, Germany
Email: labudde@hs-mittweida.de

***Abstract*—Major incidents can disturb the state of balance of a society and it is important to increase the resilience of the society against such disturbances. There are different causes for major incidents, one of which are groups of individuals, for example at demonstrations. The ideal way to handle such events would be to prevent them, or at least provide information to ensure the appropriate security services are prepared. Nowadays, a lot of communication, even criminal, takes place in social networks, which, hence, provide the ideal ground to gain the necessary information, by monitoring such groups. In the present paper, we propose an application framework for knowledge-based social network monitoring. The ultimate goal is the prediction of short-term activities, as well as the long-term development of potentially dangerous groups, based on sentiment and topic analysis and the identification of opinion-leaders. Here, we present the first steps to reach this goal, which include the assessment of the risk for a major incident caused by a group of individuals based on the sentiment in the social network groups and the topics discussed.**

Keywords–forensic; opinion-leader; topic mining; expert system; text analysis; classification; sentiment analysis

I. INTRODUCTION

The representation and communication of individuals, companies and organizations, using the Internet, especially social networks, has become the standard in our society. Even though social networks are successful and have progressed throughout these past years, they have also contributed to the formation of new criminal energy. As already mentioned in [1], in particular, the provision of an infrastructure for rapid communication and the possibility to exchange ideas, pictures etc. in private and protected environments, which are difficult to control by investigators - if at all - enables radical or extreme political groups, criminal gangs or terrorist organizations to use Social Networks as a tool to plan, appoint and execute criminal offenses. These groups often use large-scale events with a high degree of group dynamics to promote their ideas. Events, such as sporting events, demonstrations or festivals, cause high expenses on security personnel. The inherent group dynamics cause a great uncertainty and unpredictability concerning the development of such events and make it difficult to estimate how much security personnel is needed. For example, in 2014 the police officers spent more than two million working hours just on securing soccer games in Germany [2]. Targeted and automated monitoring of social networks, taking into account the applicable legal provisions, can particularly support strategic security planning as well as the development of effective prevention strategies. As a positive side effect, the subjective sense of security of the users is strengthened. Authorities of the federal office for the protection of the constitution, as well as intelligence services, are aware of the importance of social networks as a source for important information and increasingly focus on extracting and analyzing this information. However, at this point the extraction and evaluation of the information is done manually. Taking into account the increasing number of users worldwide – currently, for example, approximately 40 % of the population worldwide uses social networks – it has to be noticed that there is an enormous amount of potential profiles or communication to be monitored. This demonstrates the need for an automated solution that is capable of handling this amount of data and the resulting complexity.

Consequently, the design of an application framework, namely Social Network Analyzer (SoNA), for monitoring groups and organizations in Social Networks as key elements of critical events is presented to assist decision-makers. A prototype implementing parts of this framework for monitoring publicly accessible Facebook data is discussed.

The paper is segmented in six sections. The first two sections following the introduction discuss the concept of predictive policing as well as give a short overview about the language characteristics in Social Networks. These sections are followed by an outline of the framework, which is still under development, including how the dangerous militant profiles can be selected, how the risk of an event can be assessed and the opinion-leaders can be identified. In Section V, a prototypical implementation including its architecture and currently available features is presented. Finally, the paper ends with a conclusion, also discussing the progress of the work and its future development.

II. PREDICTIVE POLICING AS A TOOL FOR RESILIENCE ENGINEERING

A major incident includes a great number of casualties and/or severe property damage [3]. At large-scale events, such

as described above, there is always a possibility for a major catastrophic event to happen. However, whether or not it will happen is usually difficult to predict. Resilience is the ability of a socio-ecological system to recover from disturbances, for example a major catastrophic event, and retain or regain its identity, functions, structures and its ability to respond [4]. In a study about resilience the German Academy of Technical Sciences (acatech) developed a resilience cycle based on the Social Resilience Cycle by Edwards [5], which includes the following five stages: prepare, prevent, protect, respond and recover [6]. In order to return quickly to the defined secure state of balance [7] it becomes necessary to apply resilience engineering [8] in the sense of a technical support system, which allows to anticipate the disaster situation [6]. Crime that arises from dynamic groups at large-scale events as well as organized and especially political motivated crime regularly disturb the state of balance. Information gained from monitoring activities of such groups in the Internet and especially Social Networks can be used to predict the probability of such catastrophic events beforehand. Accordingly, the National Institute of Justice in the USA defined Predictive Policing as follows:

> "Predictive policing, in essence, is taking data from disparate sources, analyzing them and then using the results to anticipate, prevent and respond more effectively to future crime." [9]

The knowledge gained from the monitoring of suspicious groups in Social Networks directly contributes to an increase in resilience in the stages Prepare and Prevent of the resilience cycle [5] [10]. Therefore, the development of an automated solution to monitor Social Networks is an important step of resilience engineering.

III. Characteristics of Social Media Language

While the language used in chat rooms is one of the most researched topics [11], language used on the social media site Facebook seems to be one of the least researched, which is evident in the small amount of literature covering that topic [12]. Zappavigna [12] suggests that one reason might be the combination of several genres on one social media site, making the analysis very complex. Even though it is impossible to generalize the language found on the Internet [11], studies about language use for example in chat rooms or on microblogging sites, combined with the scarce literature covering some linguistic aspects on Facebook gave a starting point for an analysis. The focus of this paper is on posts and comments and, therefore, excludes messages written on the instant messenger.

In order to get a first impression of the language used in Facebook groups, a small corpus was created using posts and comments from different Facebook groups, relevant to the application of SoNA (see Table I). The structure of a "conversation" in a Facebook group is very different to the structure for example of a chat conversation. The starting point for a "conversation" on a group wall on Facebook is always a post, often written by the group itself. Afterwards, users can write a comment or reply to an already existing comment. In comparison to a chat conversation the user is not expected to write a comment immediately after a post was posted or write a reply to a comment from another user. In fact, they do not have to reply at all. This leads to the fact that "conversations" in Facebook groups are not almost-synchronous as in a chat conversation, yet clearly asynchronous [13] [14]. Therefore, it might be questioned whether to talk about "conversations" at all. Nonetheless, whenever users start a discussion on a group wall and reply to each other's comments within minutes, these conversations look very similar to chat messages. Overall, this "conversation" structure on Facebook leads to a highly complex way of communication, which makes the analysis of the language used and the meaning created difficult.

Furthermore, the wall on Facebook allows the users to include multimodal communication, by posting pictures or videos, either with a comment or with words included for example in the picture. Additionally, often posts include references to other websites or users simply repost a post from someone else. Another aspect that makes the automated analysis of meaning difficult is the language itself. Characteristics taken from studies on other Internet-based communication were used as features in an annotation with the UAM corpus tool of the small corpus mentioned above [11] [15] [16] [17] [18]. The results show clearly that there seems to be a difference between posts and comments. For example, orality, especially colloquial language, typing errors and lower case spelling of nouns seem to be more common in comments. In comparison, hashtags seemed to be used more often in posts than in comments. Furthermore, comments and posts can be distinguished by their length. While the length of posts varies between zero words (e. g., pictures) up to 892 words, the length of comments varies from 1 word up to 92 words. Moreover, these numbers show that in comments one can often find incomplete sentences. Even though, it seems that the typical features found in chat conversations are not used as often in comments and posts on Facebook, they are still present and create a challenge for the automated analysis used in SoNA. Especially, emoticons make the analysis of meaning difficult, because the way in which they are used to create meaning is complex and they can also be used to create irony [19]. This is why, so far, the sentiment analysis used in SoNA is based on word and not sentence level.

TABLE I: Summary of the corpus created under this work including different types of Facebook groups.

		subcorpus posts		subcorpus comments	
type	# groups	posts	words	comments	words
right-winged	5	46	4539	97	1559
left-winged	5	48	5003	94	1618
soccer ultras	2	20	1211	40	323
total	12	114	10753	231	3500

IV. Outline of a Framework

The analysis of social networks from the point of view of security policy pursues two main objectives. The first one is the identification and estimation of potential dangers, including their scope and location. The second one is to enable security forces to plan in the long-term. In order to do so, it is of special interest how a group is developing in terms of their size growth, their orientation or radicalization and the increase

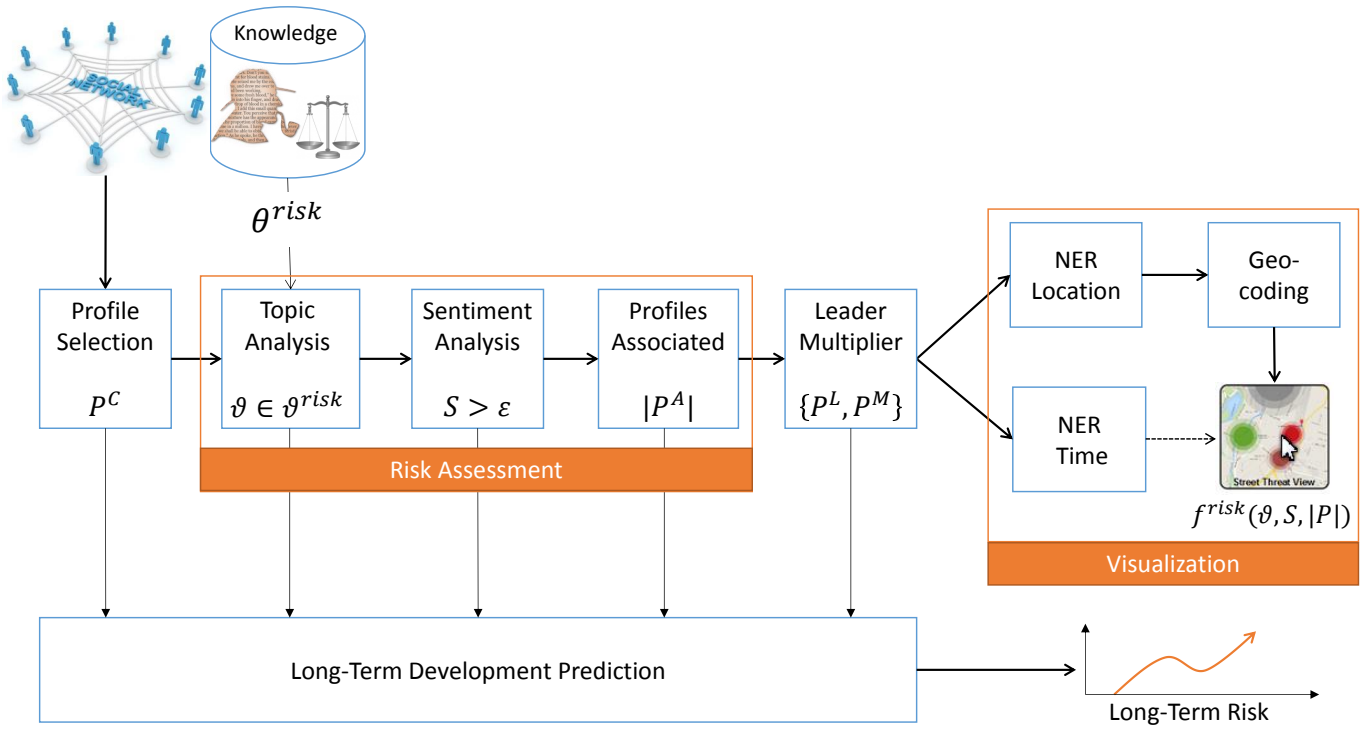

Figure 1: The proposed process chain for monitoring social networks.

in their propensity to violence. This section discusses basic concepts of a framework that addresses these tasks.

The proposed framework allows decision-makers of security forces, for example in the police's management and control centers, to identify and predict areas with high levels of crime. As a result, it is possible to deploy forces more efficiently depending on the specific situation. Thus, if, for example during a debate about the policy regarding refugees on publicly accessible pages of a social network, users loudly advocate arson attacks on refugee homes, decision-makers can now put security forces and specialized investigators on standby. If, on the other hand, before a soccer games, violent fans or fans in general do not seem to plan any riots, it may be sufficient to return to the minimum number of necessary staff to secure the event.

Another goal is to predict the long-term development of potentially violent groups. Such a prediction may include statements about the expected development of their membership, but also evidence of a possible increase in radicalization in the future. With this information, executives will be able to plan resources and make infrastructural decisions in the long term. If, for example, a district becomes, in the future, a point of attack for various, growing and violent political groups, due to certain circumstances, this information could lead to the construction of an additional police station or the expansion of the forces of an existing one. The development of a framework for the automated analysis of data from social networks with the aim of more effective crime prevention and defense in the long and short term, makes it an application from the field of predictive policing as defined in Section II.

In particular, the following tasks must be addressed by the framework:

1) selection of potential profiles of dangerous militants,
2) assessment of the probability that the danger occurs,
3) determination of location and time of risks.

In order to meet the special needs and challenges of forensics, especially with regard to the dynamics of language in social networks, it is necessary to resort to expert knowledge. This knowledge can be represented in the form of a Forensic Topic Map (FoTM) as explained in detail in [20]. In particular, abstract threats are modeled here, which form the basis for the assessment and evaluation of the communication content. Figure 1 shows the entire process chain for the proposed framework. All process steps except for the long-term prediction, which will be covered in future work, are explained in more detail in the following subsections.

A. *Selection of dangerous militants profiles*

The selection of so-called dangerous militants profiles ensures that profiles are not arbitrary selected and is thus essential to the observance of data privacy protection regulations. Furthermore, it focuses the monitoring on those profiles and thus regulates the limitation of the analysis effort. Even though the monitoring is limited to public profiles, and therefore all information publicly available, it is important not to violate the individual feeling of freedom, especially the freedom of speech as regulated in the legal framework of the respective legislature. The concept of the potential attacker was defined by a German working group, consisting of the heads from

the State Offices of criminal investigations and the Federal Criminal Police Office, for the scope of German law as follows:

> "A dangerous militant is a person in whom certain facts justify the assumption that they will commit politically motivated offenses of considerable importance..." [21, translated by H. S.]

The extent to which this definition can be extended to other areas of organized crime and gangs, without a political motivation, remains to be legally clarified. Based on that concept, a dangerous militant profile can now be defined as follows:

> A dangerous militant profile is the profile of a dangerous militant in a social network. All profiles associated with this profile are part of the extended dangerous militant profile.

Traditionally, the selection of profiles to be monitored has been carried out manually. Appropriate candidates are selected, for example through research on the Internet or other investigations. In this way, however, new or short-term profiles are hardly detected. Here, automated approaches can help in the long-term.

For example, the task of automatically identifying a dangerous militant profile, associated with a certain crime area, given a group of profiles can be interpreted as a classification task. Let P be the set of all profiles of a particular social network, and R the set of risk classes, corresponding to an area of crime. Then the selection of potential militants profiles is a surjective mapping $f : R \to P$. An overview of classification techniques (supervised learning methods) is given, for example, in [22] [23]. However, as already emphasized by [24], a large amount of training data is needed to train classifiers with sufficient accuracy. This problem can be addressed, for example, by the use of semi-supervised learning methods, such as self-training or co-training. An overview of methods is described, e. g., in [25]. Whichever method is chosen, the performance depends on the choice of appropriate features. These should generally have sufficient discrimination power and should be as independent as possible.

Particularly in the context of social networks, the use of techniques for recommender machines is often used (push-mode) instead of classification (pull mode). Typically, such systems use Collaborative Filtering [26] [27], Content-based Filtering [28] [27], or a combination of both. In recent years, a whole series of studies have been devoted to the creation of friendships in social networks using these classic approaches [29] [30]. More recent approaches are based on social graphs [31] [32] or semantic analyzes, especially LDA, which attempt to produce recommendations based on lifestyles [33] [34] [35] [36]. However, the inclusion of structured data is more reliable than the analysis of latent topics and is therefore more suitable for classifying threats. Naruchitparames et al. presented a recommender system based on genetic algorithms [37]. As a feature (social genome), they propose the following Facebook feature:

- common friends,
- location,
- age range,
- common interests (likes and music),
- photo tags,
- events,
- groups,
- movies,
- education,
- religious and political attitude.

Manca et al. criticize earlier approaches because they do not take into account a mutual interest which is, however, necessary for friendship. They suggest a similarity-based recommender as a basis for friendships using so-called Social Bookmarks, i. e., shared bookmarks on the Internet [38]. Tags of shared images are the basis for the recommender system proposed by Cheung et al. and are another interesting feature to generate recommendations in social networks [39]. In a similar manner, a general classifier can be trained based on the profiles of known dangerous militants or offenders. For example, by means of corresponding known profiles, a classifier or a recommendation system could be implemented for detecting profiles of the hooligan scene or radical political groupings. Adapting this approach, a classifier can be trained in the sense of supervised learning, which can automatically detect such dangerous militants profiles. We can use a social feature vector $\vec{f^s}$ for each profile (see Equation (1)) as a basis for the computational task of the classification and recommendation of dangerous militants.

$$\vec{f^s} = \begin{pmatrix} friends \\ location \\ age \\ interests \\ \cdots \end{pmatrix} \tag{1}$$

Considering this as a binary classification task, we need to assign each profile $\vec{f^s}$ either to the class of potential dangerous militants profiles or not. Assuming the features f_i^s are independent, we can use the Bayes theorem for computation (see 2).

$$\hat{y} = \arg \max_{c_i \in \{0,1\}} p(c_i) \prod_{j=0}^{|\vec{f^s}|} p(f_j^s | c_i) \tag{2}$$

Although we know that this assumption is not true, experiences have shown that this approach still produces good results [40]. In general, supervised approaches need a sufficiently large set of training examples which is a problem in many cases. To overcome this, we can use a bootstrapping approach, as shown in [41].

B. Assessment of the risk of dangers

After the potential dangerous militants profiles have been selected, the content analysis of the communication takes

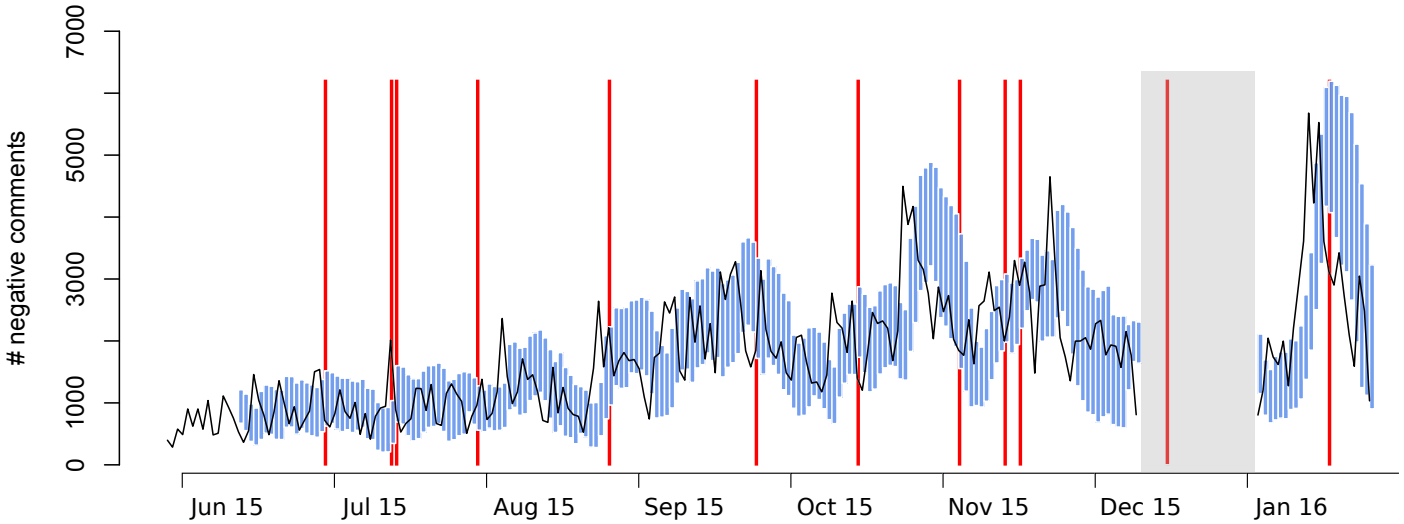

Figure 2: Results of a short-term study on the development of sentiment on the Facebook page of Pegida e. V. between June 2015 and January 2016. The blue areas mark the 95 %−prediction interval. Red lines denote actual incidents during this period of time. The gray area marks a period with missing data.

place. This step is necessary to determine whether the extraction of further information is necessary to elucidate various modalities (location, time, participants, etc.) of possible events.

A prerequisite for the assessment of the probability that the danger occurs is once again the experience-based knowledge of the investigator, which must be available for each individual risk type, for example, in the FoTM as discussed in [20] [41].

After defining the risk classes $risk_1, ..., risk_n$ which should be monitored, the explicit definition of the corresponding danger topics is made: $\Theta^{risk} = \vartheta_{risk_1}, ..., \vartheta_{risk_n}$. A risk class describes the amount of all offenses belonging to a defined group, for example, left or right-winged politically motivated crimes. A risk topic includes all the terms and associations that characterize such a risk class. Afterwards, the selection of potential or known dangerous militants profiles leads to a set of candidate profiles for each risk class $P_i^c = p_{i1}, ..., p_{ik} \in P, i = 1, ..., n$ from the set of available profiles of the investigated social network is carried out taking into account a particular risk class to limit the scope of observation and analysis. Subsequently, the topics $\Theta^{com} = \vartheta_{com_1}, ..., \vartheta_{com_n}$ of the communication between these profiles must be extracted and it must be then analyzed to what extend they overlap with the risk topics. Afterwards, they are evaluated. In the simplest case the overlap can be represented binarily as shown in Equation (3).

$$f(\Theta^{com}) = \begin{cases} 1 & if\ \Theta^{com} \cap \Theta^{risk} \neq \emptyset \\ 0 & otherwise \end{cases} \tag{3}$$

In order to quantify the degree of correspondence of Θ^{com} and Θ^{risk}, a corresponding metric is needed to compare probability distributions over the terms t of a topic. Niekler and Jähnichen examined the suitability of the Jensen-Shannon divergence, the cosine similarity, and the dice coefficient as a measurement of similarity for various topics [42]. As a result, it was found that the best results were obtained on the basis of the cosine similarity $sim(\vartheta_{com}, \vartheta_{risk})$. Adapted to the present application, $sim(\vartheta_{com}, \vartheta_{risk})$ is thus defined as:

$$sim(\vartheta_{com}, \vartheta_{risk}) = \frac{\vartheta_{com} \cdot \vartheta_{risk}}{\|\vartheta_{com}\| \|\vartheta_{risk}\|} \tag{4}$$

If $f(\Theta^{com}) = 1$, i. e., $\exists \vartheta_{com_i} | \vartheta_{com_i} \in \Theta^{risk}$, the analysis of the sentiment S in the network is carried out. Approaches are found in the literature, especially for Twitter messages [43] [44]. In principle, these approaches can also be applied to other social networks such as Facebook. If the sentiment exceeds a threshold value ε, an increased risk can be assumed.

To evaluate this hypothesis the communication on the Facebook page of "Pegida e. V." (a mostly right-winged organization in Germany) was analyzed over a period of eight months, between June 2015 and January 2016. The extracted textual communication data was divided into individual sentences (tokenization). Subsequently, one out of three polarity classes pol: positive $(+)$, negative $(-)$ or neutral (0) was assigned to each sentence s using a probabilistic language model. Equations (5) and (6) show the associated likelihood function and the derived scoring function.

$$\log_2 P(s, pol) = \log_2 P(s|pol) + \log_2 P(pol) \tag{5}$$

$$\text{score}(s, pol) = \frac{\log_2 P(s, pol)}{|s| + 2} \tag{6}$$

The polarity class with the highest score is assigned to the respective sentence. The "Multi-Domain Sentiment Encyclopedia for the German Language", which was developed at the Darmstadt University of Applied Sciences, formed the basis for the training. It contains extracted mood-bearing terms from the MLSA-A corpus [45], the pressrelations dataset [46], and the

"German Sentiment Vocabulary" (SentiWS) [47], all annotated with the average polarity values in the range $[-1, 1]$.

The sentiment of a message $m = \{s_1, ..., s_n\}$ (post, comment) is decided in the simplest case by the number of its positive sentences $s+$ and/or negative sentences $s-$. The sentiment that dominates the constituent sentences also determines the sentiment of the whole message (see Equation (7)). In case of equality, the message is considered to be neutral m^0.

$$\mathrm{S}(m) = \begin{cases} m^+ & \text{if } |s^+| > |s^-| \\ m^- & \text{if } |s^+| < |s^-| \\ m^0 & \text{otherwise} \end{cases} \tag{7}$$

This approach, of course, is only a rough estimate of the sentiment, since it does not take into account the connection between meaning (semantic) and sentiment of a sentence. The accuracy, however, appears sufficient for a first check of the hypothesis, since the messages themselves were filtered in advance by the topic analysis. The results are presented in a histogram (see Fig. 2), with only negative messages (comments) m^- being taken into account.

Comparing the development of the sentiment of the comments in the network with the events during this period (marked by red lines), it was found that there is a possible correlation between these two. For example, on January 11th, 2016 serious riots lead by the right-winged scene occurred during the demonstration of the sister organization Legida e. kV. in Leipzig (Germany). Members in the Pegida network were also encouraged to attend this event. Similar to most of the cases, it can be clearly seen that the peak of negative communication is situated immediately before the incident. The sudden reduction in conversations at the time of the event can be explained by the active participation of the members in the event. The 95%-prediction interval (blue lines) supports the assumption that incidents mostly occur after a local or global peak.

Even if this short study is not considered representative and a random correlation between the occurrence of the incident and the discussion in the network cannot be ruled out, it still shows the potential of the presented approach. At this point, additional long-term studies with larger data sets considering different networks are necessary.

C. Detection of Leaders and Multipliers

Leaders and multiplier in the context of the intended analysis of social networks are individuals, who exert a significant amount of influence on the opinion and sentiment of other users of the network through their actions. In social sciences the term 'opinion leader' was introduced before 1957 by Katz and Lazarsfeld's research on diffusion theory [48]. Their proposed two-step flow model (see Fig. 3) retains validity in the digital age, especially in the context of social media.

Katz et al. assume that information disseminated in the Social Network is received, strengthened and enriched by opinion-leaders L_i in their social environment. Since opinion-leadership is strongly knowledge-driven and thus topic-dependent, this model must be supplemented by various thematically limited opinion-leaders L_{ϑ_i}. Each individual is then influenced by a variety of heterogeneous opinion leaders in his opinion as illustrated in Fig. 3. This signifies, that the opinion of an individual is mostly formed by its social environment. In 1962, Rogers references these ideas and defines opinion leader as follows:

> "Opinion leadership is the degree to which an individual is able to influence informally other individuals' attitudes or overt behavior in a desired way with relative frequency." [49, p. 331]

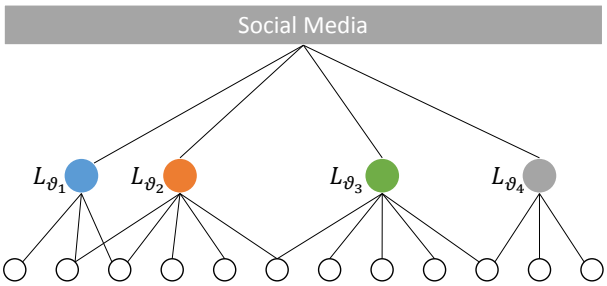

Figure 3: Extended two-step flow model adapted from [48]. Information is spread throughout social media. Individuals with a high level of competence at strategic local positions receive and amplify this information according to their competence (opinion-leader $L_{\vartheta_1}, ..., L_{\vartheta_4}$) and spread it to its followers and friends. This means, each individual's opinion is influenced by different opinion leaders depending on the topic (different colors) discussed in the network.

For the present study, the most important question to answer is what influence means, or rather how to identify an opinion leader or how the influencer can be distinguished from those being influenced. Katz defined the following features [48]:

1) personification of certain values
2) competence
3) strategic social location

One approach to identify opinion-leaders is to extract and analyze the content of nodes and edges of networks to mine leadership features. For instance, the sentiment of communication pieces can be analyzed to detect the influence of their authors, as shown by Huang et. al., who aim to detect the most influential comments in a network this way [50]. Another strategy is to perform topic mining to categorize content and detect opinion-leaders for each topic individually, as opinion-leadership is context-dependent [48] [51]. For this purpose, Latent Dirichlet Allocation (LDA) [52] can be used, as seen in the work of [53]. We considered the implementation of content-based methods problematic, as texts in social networks lack correct spelling and formal structure, which impairs such methods' performance.

Another approach to identifying leaders is to analyze the flow of information in a network. By monitoring how the interaction of actors evolves over time, one can identify patterns and individuals of significance within them. To achieve this, some model of information propagation is required, such as the

Markov processes used by [54] and the probabilistic models proposed by [55]. These interaction-based methods consider both topological features and their dynamics over time. However, the latter are not yet considered by our framework and are reserved for future developments.

We utilized methods, which are solely based on a network's topology, therefore, consider features, such as node degree, neighborhood distances and clusters, to identify opinion leaders. One implementation of this is the calculation of node centrality. The underlying assumption is, that the more influence an individual gains, the more central it is in its network. Which centrality measure is most suitable is dependent on the application domain. We judged eigenvector centrality to be most adequate, specifically Google's PageRank algorithm [56], which functions in a similar fashion. It recursively assigns a rank R to each node A, based on the rank of the nodes of its incoming edges T_i and its total number of links C_i. The value of an edge is strongly dependent on the score of its originator (see Equation (8)).

$$R(A) = \frac{1-d}{N} + d\sum_{i=1}^{n} \frac{R(T_1)}{C(T_1)}, 0 \leq d \leq 1 \tag{8}$$

With the damping factor d, normalized over the number of all nodes of the network N, a part of the resulting rank can be subtracted and distributed to all nodes. The application of PageRank for the purposes of opinion leader detection has seen merely moderate success [57] [58]. With LeaderRank, Lü et al. advocate further development and optimization of this algorithm for social networks, and have achieved surprisingly good results [59]. Users are considered as nodes and directed edges as relationships between opinion leaders and users. All users are also bidirectionally connected to a ground node. At time step t_0, all nodes receive the score $s_i(0) = 1$ except for the ground node initialized with $s_g(0) = 0$. Equation (9) describes the process of probability flow through the network, where $s_i(t)$ indicates the LeaderRank score of a node i at time step t.

$$s_i(t+1) = \sum_{j=1}^{N+1} \frac{a_{ji}}{k_j^{out}} s_j(t) \tag{9}$$

Depending on whether or not there exists a directed edge from node i to node j, the value 0 or 1 is assigned to a_{ij}. k_i^{out} describes the number of outgoing edges of a node. The final score is obtained as the score of the respective node at the convergence time t_c and the base node score at the same time, as shown in Equation (10).

$$S_i = s_i(t_c) + \frac{s_g(t_c)}{N} \tag{10}$$

The advantage of this algorithm compared to PageRank is that the convergence is faster and above all that nodes, that spread information faster and further, can be found. In later work, for example, by introducing a weighting factor, as in [60] or [61], susceptibility to noisy data has been further reduced and the ability to find influential distributors (hubs) of information has been added.

However, there might be cases in which LeaderRank would assign high scores to individuals, which are not relevant for the present application. When a user attained a significant audience, while also actively following many opinion leaders, we argue that their influence is based on their activity in the network and not their opinion, as their presence makes them more likely to be followed. We propose an approach to eliminate such hybrid leaders from the top ranks, which punishes the LeaderRank score $LR(\vartheta_i)$ of users with many interactions in the network, meaning, those users who follow many leaders. This way the top ranked users are pure leaders, whose influence is purely based on their opinions.

$$PSC(L_{\vartheta_i}) = \frac{LR(L_{\vartheta_i})}{1 + \frac{k_i^{out}}{k_{total}^{out}} * LR_{total}} \tag{11}$$

One way to calculate the PureScore of a particular topic-specific opinion leader $PSC(L_{\vartheta_i})$ is shown in Equation (11). The PureScore of a certain topic-specific opinion-leader is calculated by dividing its original LeaderRank score $LR(L_{\vartheta_i})$ by a percentage of the maximum score (equal to the number of users) defined by the node's share of network activity, k^{out} being the number of outgoing links. However, this approach needs to be evaluated in later work.

D. Visualization

If, with the approach described above, a risk greater than a threshold value ε was determined, further information such as locations, times and people involved are extracted from the text and subsequently transferred to a corresponding map with the help of geographical coordinates. An additional score $f_{risk}(\vartheta, S_\vartheta, |P|_\vartheta)$ provides information about the extent of the expected risk, estimated from the risk class, the sentiment score associated with it and the number of people involved in that particular discussion. This value can, for example, be used to color the geo location on a map, corresponding to a heat-scale. The obtained result directly supports the short-term strategic planning of security forces as proposed at the very beginning of this section.

V. Prototypical Implementation

The aim of the prototype's architecture is to implement the frameworks described in [20] as well as the sections above. It was programmed with Java and built as an Eclipse Rich Client Platform (RCP). Its OSGi implementation Equinox allows for a service-oriented architecture, consisting of three tiers:

1) *Persistence*: Data is fetched from the various social-network databases and put into *EMF* models. The models are stored into a, as for now, local *EMF Store* server. Thus, the databases and the *EMF Store* server form this tier. Any annotations and meta-data are also held by the models.
2) *Logic-Tier*: This tier contains various linguistic services, e. g., for topic modeling, used for annotation and querying. The modeling service, which provides CRUD-operations (Create, Read, Update, Delete) for models in the *EMF Store* server, also resides here.

Furthermore, all data retrieval services, which communicate with corresponding social-network APIs, are part of this tier.

3) *Access*: The high-level services, usually directly controllable by the UI, define this tier. At the current state of the development, this is the retrieval service, which initializes data fetching from the social networks, the query-service, used for data retrieval from the *EMF Store* server and the annotation services, which use the linguistic services to enrich models.

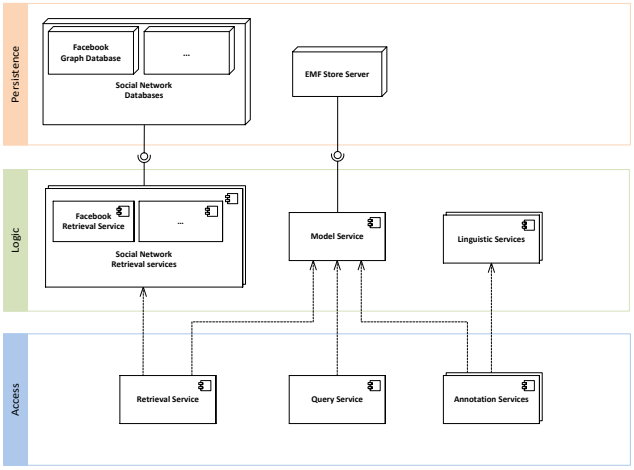

Figure 4: SoNA architecture overview with the respective services.

Figure 4 provides a visual representation of these tiers. When developing this application further, efforts will be made to make the prototype more closely resemble the SoNA framework described in this work. Permission services can be realized through user profiles in the *EMF Store* server.

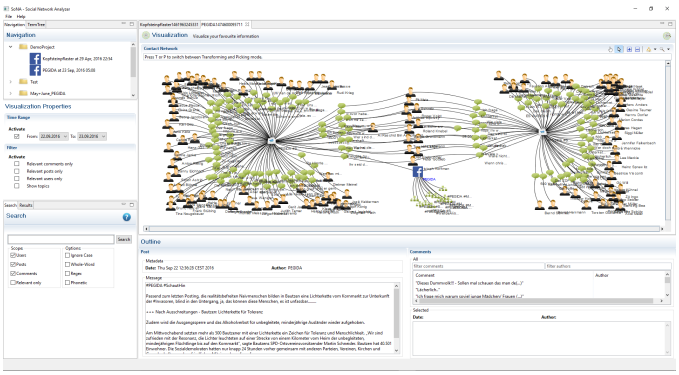

Figure 5: The user interface of the prototype. Shown is a downloaded Facebook page visualized in a graph. The Facebook logo node represents the page, which is connected to posts, which in turn are connected to their comments. The outermost nodes are users, which are associated to the comments they created.

The prototype provides a user interface to retrieve data from Social Networks, currently Facebook, and visualizes it in a 2D graph, as shown in Figure 5. Several filter options are available in order to reduce the output network to the most relevant nodes depending on the investigator's needs. Data can be retrieved from a specific part of the social network for a certain period of time or a certain amount of content (i. e., posts and comments), before being stored in models. Created models can then be annotated following the process chain as discussed in the former sections.

VI. Conclusion and Future Work

In this paper the theoretical framework SoNA was presented, which allows investigators of law enforcement agencies as well as intelligence services to monitor social networks in order to gather information about potentially dangerous activities. This information can support the long- and short-term planning of the deployment of security forces. It was shown that the knowledge gained by applying this framework can directly increase the resilience of a society in the first two stages of the resilience cycle. Furthermore, given the complexity of the language used in Social Media it was necessary to apply a knowledge-based and word-based approach. In this respect, a process chain for analyzing social networks was proposed and the main steps were discussed in detail. These include the selection of dangerous militant profiles, the assessment of the risk and the detection of leaders and multipliers. We presented a prototype of the SoNA framework, which aims to implement these aspects. Future work will include the remaining steps of the process chain as well as the evaluation of the entire framework. In order to do so, it is fundamental to create an appropriate test environment in collaboration with law enforcement agencies.

References

[1] M. Spranger, F. Heinke, S. Grunert, and D. Labudde, "Towards predictive policing: Knowledge-based monitoring of social networks," in Proceedings of the Fifth International Conference on Advances in Information Mining and Management (IMMM 2015), IARIA, Ed. ThinkMind Library, 2015.

[2] ZIS, "Jahresbericht 2013/14," 2014, [Online] https://lzpd.polizei.nrw/sites/default/files/2016-12/13-14_Jahresbericht.pdf, last accessed 2017-11-29.

[3] Deutsches Institut für Normung, "Begriffe im Rettungswesen," April 2015.

[4] J. Walker and M. Cooper, "Genealogies of resilience from systems ecology to the political economy of crisis adaptation." Security Dialogue, vol. 42, no. 2, 2011, pp. 143–160.

[5] C. Edwards, Resilient Nation. London: Demos, 2009.

[6] K. Thoma, Ed., Resilien-Tech: Resilience-by-Design; Strategie für die technologischen Zukunftsthemen, ser. acatech STUDIE. München: Utz, 2014.

[7] S. L. Pimm, The balance of nature? Ecological issues in the conservation of species and communities, 2nd ed. Chicago u.a.: Univ. of Chicago Press, 1992.

[8] C. S. Holling, "Engineering resilience versus ecological resilience," Engineering within ecological constraints, 1996, pp. 31–44.

[9] B. Pearsall, "Predictive policing: The future of law enforcement?" NIJ Journal, no. 266, 2009.

[10] K. Thoma, "Resilien-Tech. Resilience-by-Design: Strategie für die technologischen Zukunftsthemen," acatech Studie, 2014.

[11] C. Dürscheid, F. Wagner, and S. Brommer, Wie Jugendliche schreiben: Schreibkompetenz und neue Medien, ser. Linguistik - Impulse & Tendenzen. Berlin u.a.: de Gruyter, 2010, vol. 41.

[12] M. Zappavigna, Discourse of Twitter and social media: [how we use language to create affiliation on the web], ser. Continuum discourse series. London u.a.: Continuum Publ, 2012.

[13] C. Dürscheid, "Medienkommunikation im Kontinuum von Mündlichkeit und Schriftlichkeit. Theoretische und empirische Probleme," Zeitschrift für Angewandte Linguistik, vol. 38, 2003, pp. 37–56.

[14] S. Hintze, Emotionalitätsmarker in Kommentaren auf der PEGIDA-Facebook-Seite, ser. Networx, 2015, vol. 71.

[15] M. Beißwenger, Kommunikation in virtuellen Welten: Sprache, Text und Wirklichkeit ; eine Untersuchung zur Konzeptionalität von Kommunikationsvollzügen und zur textuellen Konstruktion von Welt in synchroner Internet-Kommunikation, exemplifiziert am Beispiel eines Webchats. Stuttgart: Ibidem-Verl., 2000.

[16] K. Luckhardt, "Stilanalysen zur Chat-Kommunikation: Eine korpusgestützte Untersuchung am Beispiel eines medialen Chats," Ph.D. dissertation, TU Dortmund, 2009.

[17] M. O'Donnell, "The UAM CorpusTool: Software for corpus annotation and exploration," in Bretones Callejas (Hg.) 2009 – Applied linguistics now, pp. 1433–1447.

[18] C.-V. Schnitzer, "Linguistische Aspekte der Kommunikation in den neueren elektronischen Medien SMS-E-Mail-Facebook," Doktor, Ludwig-Maximilians-Universität, München, 2012.

[19] E. Dresner and S. C. Herring, "Functions of the Nonverbal in CMC: Emoticons and Illocutionary Force," Communication Theory, vol. 20, no. 3, 2010, pp. 249–268.

[20] M. Spranger, S. Schildbach, F. Heinke, S. Grunert, and D. Labudde, "Semantic tools for forensics: A highly adaptable framework," in Proc. 2nd. International Conference on Advances in Information Management and Mining (IMMM). ThinkMind Library, 2012, pp. 27–31.

[21] B.-D. 16/3570, "Schriftliche Fragen mit den in der Woche vom 20. November 2006 eingegangenen Antworten der Bundesregierung," Drucksache des Deutschen Bundestages 16/3570 vom 24. November 2006, 2006.

[22] S. B. Kotsiantis, "Supervised machine learning: A review of classification techniques," in Proceedings of the 2007 Conference on Emerging Artificial Intelligence Applications in Computer Engineering: Real Word AI Systems with Applications in eHealth, HCI, Information Retrieval and Pervasive Technologies. Amsterdam, The Netherlands, The Netherlands: IOS Press, 2007, pp. 3–24.

[23] J. Han, M. Kamber, and J. Pei, Data Mining: Concepts and techniques, 3rd ed., ser. The Morgan Kaufmann series in data management systems. Waltham Mass.: Morgan Kaufmann, 2012.

[24] M. Ikonomakis, S. Kotsiantis, and V. Tampakas, "Text classification using machine learning techniques," WSEAS Transactions on Computers, vol. 4, no. 8, 2005, pp. 966–974.

[25] O. Chapelle, B. Schölkopf, and A. Zien, Semi-supervised learning, ser. Adaptive computation and machine learning, 2006.

[26] D. Goldberg, D. Nichols, B. M. Oki, and D. Terry, "Using collaborative filtering to weave an information tapestry," Communications of the ACM, vol. 35, no. 12, 1992, pp. 61–70.

[27] F. Ricci, L. Rokach, and B. Shapira, "Introduction to Recommender Systems Handbook," in Recommender Systems Handbook, F. Ricci, L. Rokach, B. Shapira, and P. B. Kantor, Eds. New York: Springer, 2011, pp. 1–35.

[28] R. van Meteren and M. van Someren, "Using content-based filtering for recommendation," in Proceedings of the Machine Learning in the New Information Age: MLnet/ECML2000 Workshop, 2000, pp. 47–56.

[29] L. Bian and H. Holtzman, "Online friend recommendation through personality matching and collaborative filtering," Proc. of UBICOMM, 2011, pp. 230–235.

[30] V. Agarwal and K. K. Bharadwaj, "A collaborative filtering framework for friends recommendation in social networks based on interaction intensity and adaptive user similarity," Social Network Analysis and Mining, vol. 3, no. 3, 2013, pp. 359–379.

[31] N. B. Silva, R. Tsang, G. D. C. Cavalcanti, and J. Tsang, "A graph-based friend recommendation system using genetic algorithm," in Evolutionary Computation (CEC), 2010 IEEE Congress on, 2010, pp. 1–7.

[32] F. Akbari, A. H. Tajfar, and A. F. Nejad, "Graph-based friend recommendation in social networks using artificial bee colony," in Dependable, Autonomic and Secure Computing (DASC), 2013 IEEE 11th International Conference on, 2013, pp. 464–468.

[33] N. M. Eklaspur and A. S. Pashupatimath, "A friend recommender system for social networks by life style extraction using probabilistic method-friendtome," International Journal of Computer Science Trends and Technology (IJCST), vol. 3, no. 3, 2015.

[34] Z. Wang, J. Liao, Q. Cao, H. Qi, and Z. Wang, "Friendbook: A semantic-based friend recommendation system for social networks," IEEE Transactions on Mobile Computing, vol. 14, no. 3, 2015, pp. 538–551.

[35] T. R. Kacchi and A. V. Deorankar, "Friend recommendation system based on lifestyles of users," in Advances in Electrical, Electronics, Information, Communication and Bio-Informatics (AEEICB), 2016 2nd International Conference on, 2016, pp. 682–685.

[36] D. M. Jadhavar and V. R. Chirchi, "Friend recommendation system for online social networks," International Journal of Computer Applications, vol. 153, no. 12, 2016.

[37] J. Naruchitparames, M. H. Güne, and S. J. Louis, "Friend recommendations in social networks using genetic algorithms and network topology," in 2011 IEEE Congress of Evolutionary Computation (CEC), 2011, pp. 2207–2214.

[38] M. Manca, L. Boratto, and S. Carta, "Producing friend recommendations in a social bookmarking system by mining users content," in The Third International Conference on Advances in Information Mining and Management (IMMM 2013), IARIA, Ed. IARIA, 2013, pp. 59–64.

[39] M. Cheung and J. She, "Bag-of-features tagging approach for a better recommendation with social big data," in Proc. 4th. International Conference on Advances in Information Mining and Management. ThinkMind Library, 2014, p. 83 to 88.

[40] I. Rish, "An empirical study of the Naïve Bayes classifier," IJCAI 2001 Workshop on Empirical Methods in Artificial Intelligence, vol. 22, no. 3, pp. 41–46.

[41] M. Spranger and D. Labudde, "Semantic tools for forensics: Approaches in forensic text analysis," in The Third International Conference on Advances in Information Mining and Management (IMMM 2013), IARIA, Ed. IARIA, 2013, pp. 97–100.

[42] A. Niekler and P. Jähnichen, "Matching Results of Latent Dirichlet Allocation for Text," in Proceedings of the 11th International Conference on Cognitive Modeling, ICCM 2012, 2012.

[43] X. Wan, "Co-training for cross-lingual sentiment classification," in Proceedings of the Joint Conference of the 47th Annual Meeting of the ACL and the 4th International Joint Conference on Natural Language Processing of the AFNLP: Volume 1-volume 1, 2009, pp. 235–243.

[44] S. M. Mohammad, S. Kiritchenko, and X. Zhu, "NRC-Canada: Building the state-of-the-art in sentiment analysis of tweets," arXiv preprint arXiv:1308.6242, 2013.

[45] S. Clematide, S. Gindl, M. Klenner, S. Petrakis, R. Remus, J. Ruppenhofer, U. Waltinger, and M. Wiegand, "MLSA-A Multi-layered Reference Corpus for German Sentiment Analysis," in LREC, 2012, pp. 3551–3556.

[46] T. Scholz, S. Conrad, and L. Hillekamps, "Opinion mining on a german corpus of a media response analysis," in International Conference on Text, Speech and Dialogue, 2012, pp. 39–46.

[47] R. Remus, U. Quasthoff, and G. Heyer, "SentiWS-A Publicly Available German-language Resource for Sentiment Analysis," in LREC, 2010.

[48] E. Katz, "The two-step flow of communication: An up-to-date report on an hypothesis," Public Opinion Quarterly, vol. 21, no. 1, Anniversary Issue Devoted to Twenty Years of Public Opinion Research, 1957, p. 61.

[49] E. M. Rogers, Diffusion of innovations. New York: The Free Press, 1962.

[50] B. Huang, G. Yu, and H. R. Karimi, "The finding and dynamic detection of opinion leaders in social network," Mathematical Problems in Engineering, vol. 2014, 2014, pp. 1–7.

[51] P. Parau, C. Lemnaru, M. Dinsoreanu, and R. Potolea, "Opinion leader detection." in Sentiment analysis in social networks, F. A. Pozzi, E. Fersini, E. Messina, and B. Liu, Eds., 2016, pp. 157–170.

[52] D. M. Blei, A. Y. Ng, and M. I. Jordan, "Latent Dirichlet Allocation," J. Mach. Learn. Res., vol. 3, 2003, pp. 993–1022.

[53] X. Song, Y. Chi, K. Hino, and B. Tseng, "Identifying opinion leaders in the blogosphere," in Proceedings of the sixteenth ACM conference on Conference on information and knowledge management - CIKM '07, M. J. Silva, A. O. Falcão, A. A. F. Laender, R. Baeza-Yates, D. L. McGuinness, B. Olstad, and Ø. H. Olsen, Eds. New York, New York, USA: ACM Press, 2007, p. 971.

[54] B. Amor, S. Vuik, R. Callahan, A. Darzi, S. N. Yaliraki, and M. Barahona, "Community detection and role identification in directed networks: Understanding the Twitter network of the care.data debate," CoRR, vol. abs/1508.03165, 2015.

[55] M. Richardson and P. Domingos, "Mining knowledge-sharing sites for viral marketing," in Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining, Za&iuml and O. R. ane, Eds. New York, NY: ACM, 2002, p. 61.

[56] S. Brin and L. Page, "The anatomy of a large-scale hypertextual web search engine," Comput. Netw. ISDN Syst., vol. 30, no. 1-7, Apr. 1998, pp. 107–117.

[57] C. Egger, "Identifying Key Opinion Leaders in Social Networks: An Approach to use Instagram Data to Rate and Identify Key Opinion Leader for a Specific Business Field," Master Thesis, TH Köln - University of Applied Sciences, Köln, 2016.

[58] M. Z. Shafiq, M. U. Ilyas, A. X. Liu, and H. Radha, "Identifying leaders and followers in online social networks," IEEE Journal on Selected Areas in Communications, vol. 31, no. 9, 2013, pp. 618–628.

[59] L. Lü, Y.-C. Zhang, C. H. Yeung, and T. Zhou, "Leaders in social networks, the Delicious case," PLoS One, vol. 6, no. 6, 2011, p. e21202.

[60] Q. Li, T. Zhou, L. Lü, and D. Chen, "Identifying influential spreaders by weighted LeaderRank," Physica A: Statistical Mechanics and its Applications, vol. 404, no. Supplement C, 2014, pp. 47–55.

[61] Z. H. Zhang, G. P. Jiang, Y. R. Song, L. L. Xia, and Q. Chen, "An improved weighted leaderrank algorithm for identifying influential spreaders in complex networks," in 2017 IEEE International Conference on Computational Science and Engineering (CSE) and IEEE International Conference on Embedded and Ubiquitous Computing (EUC), vol. 1, 2017, pp. 748–751.