No Mayfly: Detection and Analysis of Long-term Twitter Trends

John Ziegler[1], Michael Gertz[1]

Abstract: The focus of social media is characterized by stories about short-lived breaking news. Often, such "mayflies" make it hard to keep track of more profound topics that are prevalent over a long period of time. To provide such capabilities, we present a method to detect long-term trends based on temporal networks and community evolution. Connecting those methods with trend analysis approaches allows to study the temporal development of trends, their contextual information and how they are interrelated over time, which is of great benefit compared to existing work. Results obtained from a Twitter case study are discussed in detail and evaluated based on real-world event linkage, which proves the good functionality of the proposed method.

Keywords: Social Media Analytics; Temporal Networks; Trend Analysis; Twitter Data

1 Introduction

In today's social media landscape discussed topics and attention are rapidly changing. It is hard to not get distracted by short-lived trends ("mayflies") and instead keep focused on more profound and steady topics. In this work, inspired by the slow journalism movement [Le15], we do not analyze breaking news and trends of short attention but instead, focus on long-term trends. For this, a framework to detect and analyze long-term social media trends is outlined. It builds on existing work that is adopted and extended to fit the use case requirements. These extensions include: 1. Leveraging a temporal network model to study long-term trends, 2. Pruning of less prevalent nodes based on a power law degree distribution model, 3. Temporal tracking of hashtag communities via a core of central nodes, and 4. Appropriate visualizations to analyze the temporal development of found trends. The proposed methodology is applied to the German political Twitter-sphere to analyze long-term political trends. Thereby, the network-based approach allows to intuitively represent detected trends within their semantic context. In contrast to related work, e.g., the work by Chae and Park [CP18], our analysis specifically investigates semantic shifts of detected trends over time.

Regarding the used terminology, we do not refer to "trends" as they are often used in a time series analysis setting, e.g., [CC08, pp. 27-54]. Instead, trends in our social media analysis setting do come with a semantic meaning. Asur et al. describe trends as topics that

[1] Heidelberg University, Institute of Computer Science, Im Neuenheimer Feld 205, 69120 Heidelberg, Germany; ziegler@informatik.uni-heidelberg.de, gertz@informatik.uni-heidelberg.de

doi:10.18420/BTW2023-17

"[. . .] capture the attention of a large audience [. . .]" [As11]. We follow this definition and start by taking Twitter hashtags as representatives of topics, which is in line with previous work, e.g., [As11] [BAE11]. According to Bhulai et al. [Bh12], these hashtags might also be clustered. As a result, we extend the previous definition of a "topic" and do not refer to it as a single hashtag, but as a community of hashtags. Tracking those communities of hashtags over time results in "temporal topics". A topic can be said to make up a "trend" if its popularity is large enough (cf. [As11]) and is further called a "long-term" trend if it is prevalent over a sufficiently long time period. Together, we denote them as "long-term topical trends". Further, for differentiation between short- and long-term trends, we refer to the concept of "news cycles" or rather "political information cycles" as described by Chadwick [Ch11]. These cycles describe news production processes and typically cover a time span of a few days. Topics that are discussed in the context of such short-lived media attention cycles are defined as short-term trends. In contrast, long-term trends describe topics that are prevalent in media for several weeks, months, or even years. This distinction is in line with past work, e.g., [Ha16].

This paper is structured as follows: First, in Section 2 related work is described and compared. Section 3 then covers the methodology concerning the detection of long-term trends. The proposed method is applied to a collected political Twitter dataset, and the according analysis is described and evaluated in Section 4. Finally, Section 5 gives a summary of the present framework and describes future work. Also, the source code used for the analysis steps is publicly available at the following URL: `https://github.com/jomazi/twitter-long-term-trends`.

2 Background

Most studies related to social media trend analysis focus on short-lived and mostly event-driven scenarios, e.g., [As11] and [BAE11]. Nevertheless, Chae and Park [CP18], as an example, apply topic detection to a long-term Twitter dataset and investigate trends within the corporate social responsibility domain. They study how the popularity of topics changes over time and how topics are interrelated. In contrast to their work, we focus on the political domain, specifically aiming to analyze topical shifts over time and follow a temporal network-based approach. Also related to trend analysis, Annamoradnejad and Habibi [AH19] study the trends published by Twitter itself. Thereby, they analyze the trending time as well as the trend's re-occurrence over time. Further, Majdabadi et al. [Ma20] propose a graph-based Twitter trend extraction method and do not only take hashtags but also terms into account. Still, they do not track those trends over long time periods. Similarly, the work by Khan et al. [Kh21] is dealing with the detection as well as ranking of trends based on Twitter data. Some existing work from the field of information retrieval also approaches trend-related use cases. As an example, Hashvati et al. [Ha16] propose an online method to detect trends in a user search context. Notably, they also use social network communities as trend candidates and distinguish between short- and long-term trends. Further, focused on classifying trends on Twitter, the work by Zubiaga et al. [Zu15] outlines a classification system of Twitter trends, along with methods to correctly identify a trend's category at its initial stage. For trends, they rely on the official trends shown on the Twitter platform. These trends are short-living [Tw] and, are either related to news, ongoing events, memes or commemoratives [Zu15].

3 Methodology

In the following section, the methodology underlying the detection of long-term trends is outlined. For this, we first introduce the leveraged dataset in Section 3.1, then continue by describing the temporal network-based model formalism in Section 3.2 and outline the processing of the used hashtag co-occurrence networks in Section 3.3. Finally, Sections 3.4 and 3.5 cover the detection of topics and their tracking over time, which also leads to the extraction of topical long-term trends.

3.1 Dataset

The EPINetz Twitter Politicians Dataset 2021 provides "[. . .] Twitter accounts of German parliamentarians, ministers, state secretaries, parties, and ministries on a state, federal, and European Union level for the year 2021" [Kö22]. We rely on the Twitter search API v2[2] to gather the raw tweets based on those user accounts. We collect tweets posted by the 2,449 accounts for the time range from January 2021 until July 2022 without filtering. In total, the dataset contains about 1.8 million tweets. Hashtags used in the tweets are taken as representatives of topics, which corresponds to the procedure of other works [As11] [BAE11]. We extract timestamped information about the (co-)occurrence of the hashtags from the unprocessed tweets and use them as the basis for detecting long-term topical trends.

3.2 Temporal networks

Taking the timestamped information about hashtag (co-)occurrences as described above, temporal networks are created as aggregations based on a given time window. To formally describe the temporal snapshot networks we rely on the framework of multi-slice networks as outlined by Bianconi [Bi18, pp. 106-110]. A multi-slice temporal network is a special kind of multilayer network with each layer/slice representing a temporal snapshot of the complete network. As in our case, no interactions across snapshots exist, we focus on the intralink networks only, i.e., multi-slice networks without interlinks. Such a multilayer network M is defined as a tuple, $M = (L, \mathcal{G})$. It consists of the network layers L with $|L| = l$.

[2] Twitter Developer Platform: https://developer.twitter.com/en/docs/twitter-api/tweets/search/introduction; Accessed 28-12-22

A single layer is referred to as $\ell \in L$. Additionally, $\mathcal{G}$ describes the time-ordered list of networks that are made up of the interactions within each of those layers:

$$\mathcal{G} = (G_1, G_2, \ldots, G_\ell, \ldots, G_l) \quad \text{with} \quad G_\ell = (V_\ell, E_\ell) \tag{1}$$

Each network G_ℓ consists of a set of nodes V_ℓ, which are in our case hashtags and their co-occurrences as edges E_ℓ. Given that a multi-slice network M covers the interactions within a time period T and the time-window Δt is chosen as snapshot size, e.g., one month, there are $l = T/\Delta t$ layers. Thereby, layer ℓ captures the interactions that occur in the timeframe $[(\ell - 1)\Delta t, \ell\Delta t)$. Within such a layer ℓ the degree of a node i is denoted as k_i^ℓ. Further, for the *aggregated* network $\tilde{G}$ of the multi-slice network, the temporal nature of the interactions is simply neglected and edges from all snapshots are taken into account.

3.3 Network processing

In contrast to mostly event- or breaking news-related short-term trends [Zu15], which are often represented by a single hashtag, long-term trends deal with more complex topics and can therefore be seen as communities of interrelated hashtags (see Section 1). To obtain more meaningful community networks and to further save computational costs during the community detection step (see Section 3.4), we focus on popular and highly connected hashtags. For this, less connected hashtags, i.e., with a low co-occurrence degree, are removed from the temporal networks. We take the median node degree per snapshot as a reference and remove all hashtags with a degree below this threshold from the according temporal network. An investigation of the degree distribution reveals its power law nature ($k \propto k^{-\alpha}$). Therefore, we leverage the median as defined by Newman [Ne05]:

$$k_{med} = 2^{1/\alpha - 1} k_{min} \tag{2}$$

An exemplary degree distribution is shown in Figure 1. The fitting procedure, for which the "powerlaw" package provided by Alstott et al. [ABP14] is used, reveals a power law exponent of 1.42 and according to that a median k_{med} of 5.31. In addition to the pruning step, the temporal snapshot networks are weighted. Ideally, respective edge weights reflect the semantic expressiveness of a hashtag and the strength of interrelations between hashtags. For this, we refer to Pointwise Mutual Information (PMI) [RN11]. Given that f_i^ℓ describes the frequency of occurrence of node i during the timeframe covered by layer ℓ and f_{ij}^ℓ the frequency of co-occurrence of nodes i and j, the according PMI value is defined as:

$$\mathrm{PMI}_{ij}^\ell = \ln \frac{f_{ij}^\ell}{f_i^\ell \cdot f_j^\ell} = w_{ij}^\ell \tag{3}$$

As indicated in Equation 3, those PMI values are used as co-occurrence edge weights w_{ij}^{ℓ} between hashtag i and j in layer ℓ of the temporal multi-slice network.

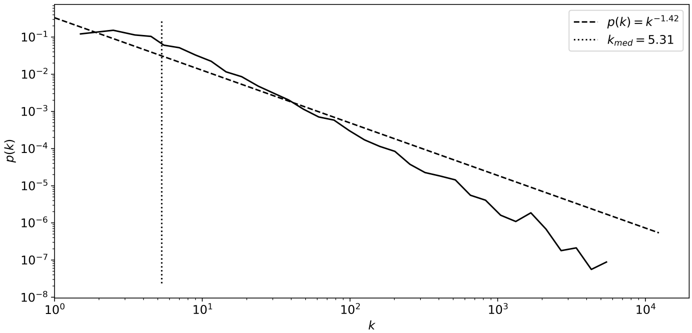


Fig. 1: Degree distribution of the January 2021 network snapshot

3.4 Detection of hashtag communities

Hashtags, i.e., the nodes of the temporal networks, are taken as representatives of topics [As11] [BAE11]. Further, according to Bhulai et al. [Bh12] in a comprehensive trend analysis framework related topics should be clustered. Therefore, we rely on methods developed in the field of community detection to find groups of densely interrelated hashtags. Those groups of hashtags then form a topic with all of its aspects as multiple hashtags might describe different semantic dimensions of the topic. To be precise, we leverage the Leiden community detection algorithm by Traag et al. [TWV19] and use the implementation as provided by the igraph software package [CN+06]. The community detection is applied to all layers of the temporal network described in Section 3.2.

3.5 Long-term trend detection

Of course, temporal communities of hashtags, i.e., temporal topics, as described in Section 3.4 do not yet make up a long-term topical trend. Asur et al. describe trends as topics that "[. . .] capture the attention of a large audience [. . .]" [As11], which means that trends need to reach a certain level of popularity. For this to measure, we take the accumulated count of hashtag occurrences per community and time window as trend scores. A community i in the network layer/slice ℓ is given as a subset of hashtag nodes: $C_i^{\ell} \subseteq V_{\ell}$. Together with a mapping of those nodes to their respective occurrence counts for the given layer ℓ, $o_{\ell} : V_{\ell} \to \mathbb{N}$, we define the trend scores τ as follows:

$$\tau(C_i^{\ell}) = \sum_{v \in C_i^{\ell}} o_{\ell}(v) \tag{4}$$

Those scores allow to rank detected trends by their popularity and, as an example, only the top-n trends can be investigated. *Long-term* trends, opposed to short-lived trends, need to be prevalent over a sufficiently large time span. Therefore, detected hashtag communities need to be tracked over time. In their work, Lorenz et al. [Lo17] specifically propose a method to capture the dynamics of weighted hashtag co-occurrence networks. Not only does their method allow to track communities of hashtags across subsequent time steps, but also across further distant snapshots. Considering higher-order memory, i.e., taking the networks of multiple previous snapshots into account, their approach allows to overcome issues related to temporal fluctuations and instabilities of the single-layer (static) community detection process. We built on this existing work and leverage their approach to track popular temporal hashtag communities over time, which then form long-term topical trends.

4 Analysis and Evaluation

To illustrate the long-term trend detection method described in Section 3, it is applied to the political Twitter dataset as outlined in Section 3.1. Extracted hashtag co-occurrences are aggregated into monthly snapshots. For a global description of a trend, independent of time, the aggregated network as described in Section 3.2 is leveraged. As described in Section 3.4, the Leiden algorithm [TWV19] is used for the community detection step. We use modularity as the objective function along with a resolution parameter of 1, $\beta = 0.01$ and 1000 iterations. Edge weights as outlined in Section 3.3 are taken into account. The algorithm is applied 10 times, and only the clustering that leads to the highest modularity score is taken to define the communities of hashtags, i.e., topics. Of course, due to the built-in randomness, repeated runs do not always lead to the exact same results but slight variations might occur. Per community, the induced subgraph of the 10 nodes, i.e., hashtags, with the highest PageRank scores [Pa99] is taken to represent a trend. Trend networks consist of those hashtags as nodes and their weighted interactions. As many communities contain hashtags that are either used for only a short time on social media or are very specific, we focus on the set of the 25 most central nodes, according to their PageRank, and link communities according to the similarity between those sets. For this, we leverage the method proposed by Lorenz et al. [Lo17] as described in Section 3.5. Four months are used as memory for the matching procedure to also link communities with temporal fluctuations and focus on the long-term prevalence of a trend.

In the following Section 4.1, we present analysis results covering the prevalence of trends over time, their evolution in Section 4.2, and temporal interactions in Section 4.3. Finally, results are evaluated in Section 4.4.

4.1 Prevalence of trends

As topics are tracked over time, their prevalence and popularity can be investigated with a focus on their temporal development. Not all topics might be equally prevalent at a given point in time, nor might they be occurring across all time windows. Also, the popularity of an individual topic might change significantly over time. Figure 2 shows a temporal heatmap of the trend scores as outlined in Section 3.5. Trend scores are normalized on a trend basis, meaning that a value of 1 indicates the maximum of reached popularity for an individual trend. The heatmap shows the 10 trends with the overall highest trend scores and visualizes their development over the 18 months of the entire dataset.

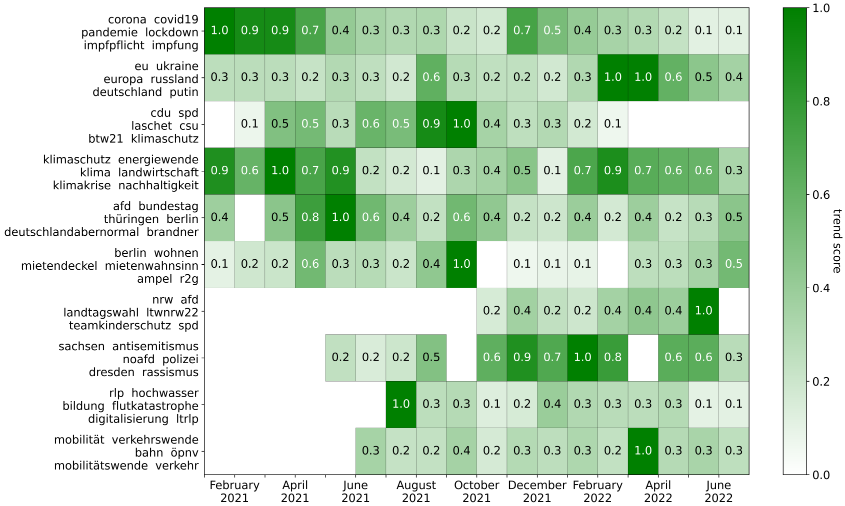
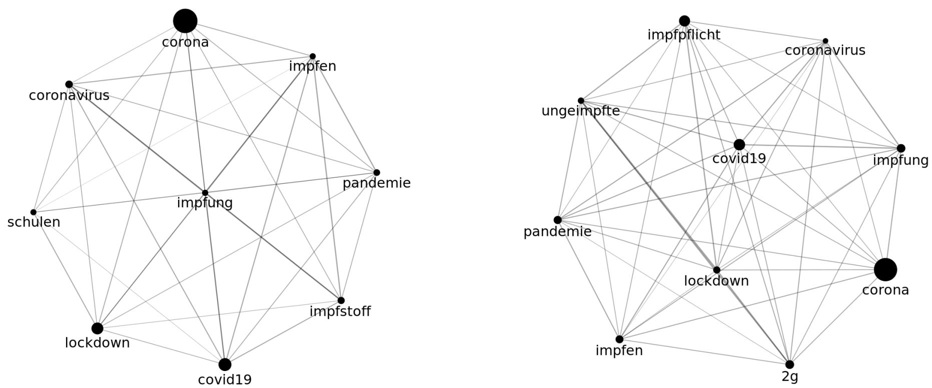


Fig. 2: Temporal heatmap of trend scores

First of all, it has to be noted that some trends, such as the one related to foreign policy and the European Union, are present across the entire time span whereas, for others, gaps in their prevalence over time become visible. That those gaps are occurring in the trend detection results confirms that the used method is indeed capable of handling temporal fluctuations. The topic is tracked over time even though it might not be detected in all intermediate snapshots. In contrast, some trends do not show gaps but are only present for a limited time span. As further described in Section 4.4, those trends are often related to some sort of event, e.g., the flood in the Ahr region. During the occurrence of that event, the trend's popularity is often at its high. All trend developments show periods of higher and lower prevalence. As an example, the COVID-19-related long-term trend is most prevalent during the spring and winter of 2021, which might be due to a more tense pandemic situation during those periods. Further, some trends do peak at approximately the same time. Of course, one cannot conclude any causality or correlation from that but at least the heatmap makes such patterns visible. Exemplary of this are the peaks of the trends related to the Russian invasion of Ukraine, which also triggered an ongoing media discussion about public transportation ("mobilität", "verkehrswende") and renewable energy ("klimaschutz", "energiewende").

4.2 Temporal evolution

Topical trends do usually not consist of only a single keyword but are instead described by multiple aspects. With the proposed trend networks those aspects, represented by hashtags, and their interrelations are intuitively visualized. More interestingly, by tracking them over time the temporal changes in the topical trends can be analyzed. As an example, see Figure 3 that shows the trend networks related to the COVID-19 pandemic, as indicated by the respective hashtags, for the two time periods of January and November 2021. For the graph layout, the igraph [CN+06] implementation of the Fruchterman and Reingold [FR91] algorithm is used. Even though some hashtags can be found in both networks, e.g., "corona" and "pandemie", other aspects and their importance change over time, e.g., "lockdown" and "impfstoff" vs. "impfpflicht" and "2g". Also, it seems as for this trend, hashtags are a lot more interrelated during November 2021 as more edges in the network show. Represented by their weighting, those edges also indicate relationships of different strengths.



(a) January 2021 (b) November 2021

Fig. 3: Trend networks related to the COVID-19 pandemic covering different time periods

4.3 Trend interrelation

Chae and Park [CP18] already highlight the importance of topic interrelations. We go in the same direction and analyze *temporal* interrelations between topics. Topics do not co-exist independently of each other, but might instead be merged over time or at least become more or less interrelated.

Figure 4 visualizes the temporal interrelations between tracked trends for the time period of February until March 2022. The veins of the alluvial diagram [RB10] represent the flow of nodes between two communities and therefore, also the interrelation between topics across time. For the most part, topics seem to be quite stable as the majority of nodes stays

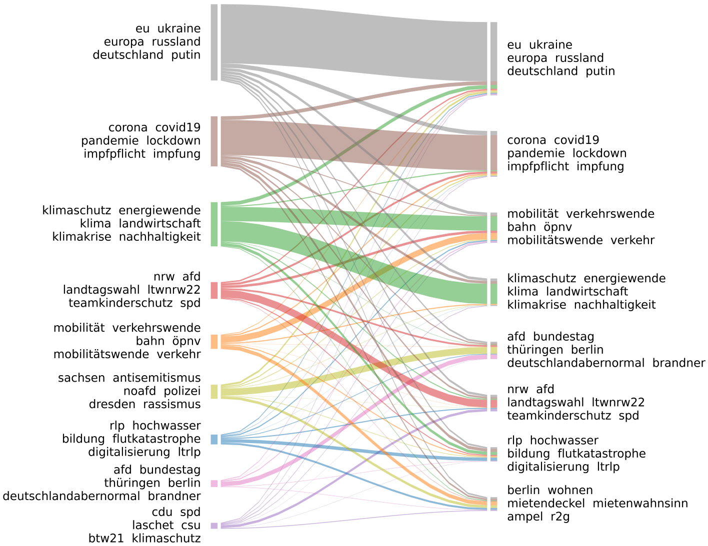


Fig. 4: Alluvial diagram visualizing the temporal interrelation of trends

within the same community. Nevertheless, some topics, e.g., the one related to climate protection, also influence multiple other ones, and nodes of these communities move to other topics. Most notably, a large portion of the climate protection topic shifts to the public transportation topic. To quantify these observations, 114 hashtags stay in the community, whereas 81 shift to the public transportation-related topic. Additionally, 21 shift to the foreign policy topic and 17 go to the one covering the Ahr flooding (see Section 4.4). Those results indicate a context switch of certain topical aspects as they become relevant for other trends as well.

4.4 Evaluation

To confirm that computed trends are actually meaningful, we leverage an event-based evaluation and manually check if detected trends are related to real-world events. For the top 10 most prevalent trends (see Figure 2), the time frame of their highest popularity is taken as prediction and related events are checked for their temporal occurrence as kind of ground truth. In a subsequent step, the trend peak and the temporal occurrence of the related event are then compared and checked for accordance.

Table 1 shows that half of the top 10 long-term trends can be related to events, such as the COVID-19 pandemic or the Russian invasion of Ukraine. Popularity peaks of these trends are in close temporal proximity to the occurrence of the related events. We argue that for

Tab. 1: Long-term trends and related events

	Hashtags	Peak	Event	Reference (accessed 28-12-22)
1	corona, covid19, pandemie, lockdown, impfpflicht, impfung	January 2021	COVID-19 pandemic (17 November 2019 – present)	`https://en.wikipedia.org/wiki/COVID-19_pandemic`
2	eu, ukraine, europa, russland, deutschland, putin	February and March 2022	2022 Russian invasion of Ukraine (24 February 2022 – present)	`https://en.wikipedia.org/wiki/2022_Russian_invasion_of_Ukraine`
3	cdu, spd, laschet, csu, btw21, klimaschutz	September 2021	2021 German federal election (26 September 2021)	`https://en.wikipedia.org/wiki/2021_German_federal_election`
4	klimaschutz, energiewende, klima, landwirtschaft, klimakrise, nachhaltigkeit	March 2021		
5	afd, bundestag, thüringen, berlin, deutschlandabernormal, brandner	May 2021		
6	berlin, wohnen, mietendeckel, mietenwahnsinn, ampel, r2g	September 2021		
7	nrw, afd, landtagswahl, ltwnrw22, teamkinderschutz, spd	May 2022	2022 North Rhine-Westphalia state election (15 May 2022)	`https://en.wikipedia.org/wiki/2022_North_Rhine-Westphalia_state_election`
8	sachsen, antisemitismus, noafd, polizei, dresden, rassismus	January 2022		
9	rlp, hochwasser, bildung, flutkatastrophe, digitalisierung, ltrlp	July 2021	Flooding of Ahr and Eifel region in Germany (15 July 2021)	`https://www.dw.com/en/flooding-in-germany-before-and-after-images-from-the-ahr-and-eifel-regions/a-58299008`
10	mobilität, verkehrswende, bahn, öpnv, mobilitätswende, verkehr	March 2022		

the other trends as well meaningful descriptions can be found, like “climate protection” for trend 4, “AfD party” for trend 5, “housing market” for trend 6, “discrimination” for trend 8 and “public transportation” for trend 10. Nevertheless, those trends are not directly linked to real-world events. Together, the event-referenced and manually labelled trends prove good functionality of our long-term trend detection method.

5 Conclusion and Future Work

This work tackles the issue of detecting long-term prevalent topics, hidden in the large volume of short-lived news media. Based on methods known from the field of temporal network analysis and community evolution, an approach to detect such long-term trends is presented. A case study based on German political Twitter data proves that actually meaningful trends are detected. For a lot of the top trends, related real-world events can be identified, as shown in Section 4.4. Future work might target more extensive evaluation procedures and additional quantitative metrics to describe the long-term evolution of trends. Also, the current trend detection approach could be extended by a more sophisticated semantic topic model.

Acknowledgements: We thank the Klaus Tschira Foundation for funding this research in the framework of the EPINetz project: `https://epinetz.de`.

References

[ABP14] Alstott, J.; Bullmore, E.; Plenz, D.: powerlaw: a Python package for analysis of heavy-tailed distributions. PloS one 9/1, e85777, 2014.

[AH19] Annamoradnejad, I.; Habibi, J.: A comprehensive analysis of twitter trending topics. In: 2019 5th International Conference on Web Research (ICWR). IEEE, pp. 22–27, 2019.

[As11] Asur, S.; Huberman, B. A.; Szabo, G.; Wang, C.: Trends in social media: Persistence and decay. In: Proceedings of the International AAAI Conference on Web and Social Media. Vol. 5. 1, pp. 434–437, 2011.

[BAE11] Budak, C.; Agrawal, D.; El Abbadi, A.: Structural trend analysis for online social networks. Proceedings of the VLDB Endowment 4/10, pp. 646–656, 2011.

[Bh12] Bhulai, S.; Kampstra, P.; Kooiman, L.; Koole, G.; Deurloo, M.; Kok, B.: Trend visualization on Twitter: what's hot and what's not? In: 1st International Conference on Data Analytics. Pp. 43–48, 2012.

[Bi18] Bianconi, G.: Multilayer networks: structure and function. Oxford university press, 2018.

[CC08] Cryer, J. D.; Chan, K.-S.: Time series analysis: with applications in R. Springer, 2008.

[Ch11] Chadwick, A.: The political information cycle in a hybrid news system: The British prime minister and the "Bullygate" affair. The International Journal of Press/Politics 16/1, pp. 3–29, 2011.

[CN+06] Csardi, G.; Nepusz, T., et al.: The igraph software package for complex network research. InterJournal, complex systems 1695/5, pp. 1–9, 2006.

[CP18] Chae, B.; Park, E.: Corporate social responsibility (CSR): A survey of topics and trends using Twitter data and topic modeling. Sustainability 10/7, p. 2231, 2018.

[FR91] Fruchterman, T. M.; Reingold, E. M.: Graph drawing by force-directed placement. Software: Practice and experience 21/11, pp. 1129–1164, 1991.

[Ha16] Hashavit, A.; Levin, R.; Guy, I.; Kutiel, G.: Effective trend detection within a dynamic search context. In: Proceedings of the 39th International ACM SIGIR conference on Research and Development in Information Retrieval. Pp. 817–820, 2016.

[Kh21] Khan, H. U.; Nasir, S.; Nasim, K.; Shabbir, D.; Mahmood, A.: Twitter trends: A ranking algorithm analysis on real time data. Expert Systems with Applications 164/, p. 113990, 2021.

[Kö22] König, T.; Schünemann, W. J.; Brand, A.; Freyberg, J.; Gertz, M.: The EPINetz Twitter Politicians Dataset 2021. A New Resource for the Study of the German Twittersphere and Its Application for the 2021 Federal Elections. Politische Vierteljahresschrift/, pp. 1–19, 2022.

[Le15] Le Masurier, M.: What is slow journalism? Journalism practice 9/2, pp. 138–152, 2015.

[Lo17] Lorenz, P.; Wolf, F.; Braun, J.; Djurdjevac Conrad, N.; Hövel, P.: Capturing the dynamics of hashtag-communities. In: International Conference on Complex Networks and their Applications. Springer, pp. 401–413, 2017.

[Ma20] Majdabadi, Z.; Sabeti, B.; Golazizian, P.; Asli, S. A. A.; Momenzadeh, O.; Fahmi, R.: Twitter Trend Extraction: A Graph-based Approach for Tweet and Hashtag Ranking, Utilizing No-Hashtag Tweets. In: Proceedings of the 12th Language Resources and Evaluation Conference. Pp. 6213–6219, 2020.

[Ne05] Newman, M. E.: Power laws, Pareto distributions and Zipf's law. Contemporary physics 46/5, pp. 323–351, 2005.

[Pa99] Page, L.; Brin, S.; Motwani, R.; Winograd, T.: The PageRank citation ranking: Bringing order to the web. Tech. rep., Stanford InfoLab, 1999.

[RB10] Rosvall, M.; Bergstrom, C. T.: Mapping change in large networks. PloS one 5/1, e8694, 2010.

[RN11] Role, F.; Nadif, M.: Handling the impact of low frequency events on co-occurrence based measures of word similarity. In: Proceedings of the international conference on Knowledge Discovery and Information Retrieval (KDIR-2011). Scitepress. Pp. 218–223, 2011.

[Tw] Twitter, Inc.: Twitter Trends FAQ – trending hashtags and topics, `https://help.twitter.com/en/using-twitter/twitter-trending-faqs`, Accessed: 28-12-22.

[TWV19] Traag, V. A.; Waltman, L.; Van Eck, N. J.: From Louvain to Leiden: guaranteeing well-connected communities. Scientific reports 9/1, pp. 1–12, 2019.

[Zu15] Zubiaga, A.; Spina, D.; Martínez, R.; Fresno, V.: Real-time classification of twitter trends. Journal of the Association for Information Science and Technology 66/3, pp. 462–473, 2015.