Visual Network Analysis of Twitter Data for Co-Organizing Conferences: Case CMAD 2013

Jari Jussila
Tampere University of Technology
jari.j.jussila@tut.fi

Jukka Huhtamäki
Tampere University of Technology
jukka.huhtamaki@tut.fi

Kaisa Henttonen
Lappeenranta University of Technology
kaisa.henttonen@lut.fi

Hannu Kärkkäinen
Tampere University of Technology
hannu.karkkainen@tut.fi

Kaisa Still
VTT Technical Research Centre of Finland
kaisa.still@vtt.fi

Abstract

The aim of this research is to explore what kinds of insights information visualization of social media data can provide for co-organizing conferences. Our paper focuses on Twitter use before, during and after conference. We present a case study based on an conference of Community Manager Appreciation Day (CMAD 2013). With the process of data-driven visual network analysis, we used Twitter data to analyse the network of conference participants and the conference's discussion topics. We were able to identify e.g. influential conference participants, most interesting presentations and discussions, similarities between interests of the conference participants. Hence, several development and information needs of conference co-organization were derived from the information visualizations, which have implications for improving the planning and co-organizing of conferences, as well as for Twitter use in conference communication.

1. Introduction

The role of conferences and events is significant in the transfer of scientific, managerial and other types of information and knowledge. The use of various conference management systems (CMS) that simplify the organization of events and conferences captures some of the interactions between conference participants as well as between them and the organizers, however, many of those interactions take place outside the traditional CMS. One rich source of information for conference and event organization is increasingly social networking services and social media at large [4,38].

Quite surprisingly, however, only 39% of event organizers use currently monitoring tools to track social media conversations and success of their organization activities [4]. The use of social media in conference co-creational activities has been found to bring advantages to all stakeholders e.g. in the form of conference marketing, better understanding of participants' needs, better conference content tailoring, more efficient information sharing, and networking [e.g. 37,32,39].

Twitter, a microblogging service, is increasingly used as a technological platform for conference co-organization, i.e. as a digital backchannel, which "constitutes a complex multidirectional discursive space in which the conference participants make notes, share resources, hold discussions and ask questions as well as establishing a clear individual online presence" [36], also before and after the conference.

For this research, we see that Twitter can be more useful in the context of conferences than, for example, more traditional group email as it enables conference participants to ask questions and get instant responses to them. Also Twitter provides possibility to tap into the collective consciousness of other conference participants and working on ideas back and forth with other participants they would not otherwise be able to connect with.

In this research, we concentrate on how users interact during the different phases of the conference.

We chose this focus because social network structure within Twitter plays crucial role in the spread of information [30]. The network characteristics provide insight in the relational features between conference participants instead of features concerning one conference participant only.

We ask: What type of insights foremost for conference co-organizers can Twitter data combined

978-1-4799-2504-9/14 $31.00 © 2014 IEEE
DOI 10.1109/HICSS.2014.190

with the possibilities of information visualization, provide of networks of people and networks of discussions before the conference, during the conference and after the conference?

As a result of this study, we identify influential conference participants, most interesting presentations and discussions, similarities between interests of the conference participants, and, most importantly, several development and information needs of conference co-organization.

More from the managerial perspective, we aim to find out how conference co-organizers could facilitate and support interaction between conference attendees on Twitter platform in practice, and thus, better serve the attendees in various ways? Further, we enquire what conference organizers should do to enable co-organization in this context. In comparison to our previous study [28] of Twitter data for co-organising organizing conferences the current study provides much more detailed information on the dynamics of networks during the lifetime of a conference whereas the previous study mainly focuses on the during conference-phase. This type of more detailed information is important for several reasons. For example, it enables us to learn how to create networks and networks structures that benefit network participants. Also, it enables us in longer run to understand how network participants can take network-modifying actions to create network structures for their benefit. [see also 3] The paper is organized as follows. We, first, review existing literature related to Twitter at conferences as well as approaches to analyze its interactions. Then we present our research method and approach of visual network analysis, applied in the case context of CMAD 2013 conference. After presenting the findings, we provide discussion and conclusions as well as arenas for further research.

2. Related research

We acknowledge that there is a lot of research on Twitter in general. Extensive reviews have also been conducted on the theme [see for example 43,1]. In their review William, Terras and Warwick [43] do not explicitly focus on analyzing the previous studies from the perspective of network characteristics and relational issues. They rather focus their review on characteristics related to message, user, technology, concept, domain, data and method. On the other hand, as a result of their literature review Aarts, Maanen, Ouboter and Schraagen [1] conclude that network characteristics and relational aspects have received less attention in the literature in comparison to actor characteristics and message characteristics.

In this paper, we are specifically interested in network characteristics and relational aspects in the context of using Twitter in conferences. We, more specifically, approach this phenomenon from the perspective of conference co-organization. This type of research still seems quite scarce to date; however, two related streams of research can be identified. First, there are studies that have focused on those who use Twitter [14,37] and the means how they use it during conferences [31,39]. The second stream concentrates on interaction on Twitter during the different phases of a conference, with some of the studies focusing on types of interaction during the conference and whether Twitter encourages participatory culture [33,39]. Furthermore, there are studies that focus on the interaction before the conference, during the conference and after the conference [16] by analyzing the number of tweets in each phase of a conference. The results of these studies show that the relative number of tweets is highest in during conference-phase.

To summarize, we still do not seem to know much about how the users interact during the different phases of the conference.

Traditionally, the role of conference participants has been relatively passive. By novel ways of integrating participants in the pre-conference planning, during-conference participation and after-conference activities of for example sharing conference-related information, conference contents and networking can be tailored to be more useful for the participants. In addition, conference marketing can be crowdsourced to conference participants and organizers alike. In recent years, conferences have adopted ways of integrating for instance social media related approaches for the purpose of the above activities, and related research has been published on the generic topic of conference co-creational activities [e.g. 19,16,24,37,15] and related benefits. For instance Grimnes [19] sums up the various needs of conference participants, such as tagging the conference content and talks, as well as browsing effectively the conference contents. Such activities, among many others, can be supported with various Web 2.0 and social media-based approaches.

2.1. Twitter and conferences

Current literature brings forth the possibilities and benefits but also the restrictions of Twitter as an approach to support conference organization, conference-related collaboration and conference content sharing and communication.

The limited message size of 140 characters is often argued to be both a strength and a weakness of Twitter messaging: it can be seen as a benefit

because it requires users to condense the main message to a very short space. This also makes easier the particularly fast circulation and retweeting of interesting messages between potentially large amounts of users. Downsides include that much information can be lost e.g. if the context of messages is not mentioned or understood properly, which can for example limit the usefulness of Twitter stream for external participants that want to follow the conference from outside [15]. An important generic challenge of Twitter is also that much of the information that is delivered through Twitter feeds in conferences, be it explicitly mentioned in tweets or implicitly otherwise present in them, remains hidden or undecipherable to common analytics approaches [41].

2.1.1 Twitter use before a conference. Before conference, Twitter seems to be mainly used most often used to market events, workshops and keynote presentations or to remind attendees to register or to take specific items with them [37]. Attendees of a conference are likely to use Twitter when they are organizing their trips i.e. they may share information about local hotels and ways to travel to the conference location. Organizers, on the other hand, aim at increasing excitement about the conference and creating a community of early adopters. [see e.g. 37] Much is not known about the dynamics of interaction in this phase of a conference. According to Ebner and Reinhardt [16], the interactions (measured by the number of tweets) remain quite low before the conference.

2.1.2. Twitter use during conference. According to Stankovic et al. [41], Twitter has lately gained significant popularity among conference and organized event participants as a means for intra-event communication. Organizers of conferences use Twitter in this stage to provide information on possible last minute changes. They may also try to engage attendees to upload pictures, share links to blog entries related to the conference themes, and customize the content of individual presentations etc. The attendees' use of Twitter during the conference has been found to depend on their personal preferences and styles; Twitter can be used to write down personal notes, to ask questions about the presentation(s) and to discuss specific topics [16,39]. When it comes to the social interactions in this phase of the conference, they have been found to increase almost dramatically [16].

2.1.3. Twitter use after conference. After the conference, the organizers often use Twitter to thank attendees for their presence. Also reflections are posted and maybe also interesting statistics are delivered. In this phase, organizers also try to gather feedback and ideas for the next conference as well as inform the attendees about possible upcoming dates. Attendees in the conference most often use Twitter to inform others about their blogs. In their blogs they may have published their reflections of the conference that may be beneficial to other attendees. The community of interest may also share links e.g. to other interesting meetings. Hashtags provide ways for staying in touch also after the conference [see e.g. 37].

2.2. Twitter and interaction

The analysis on structures of online social interaction can provide insights into human interactions on the technological platforms that are being used to enable and support them [21]. Online social interaction (referred as "vast data of backchannel conversation") by Sopan et al. [40] is already rich and fairly accessible and objective data. Thus, it is much easier for a researcher to see who said what to whom and under what contingencies looking at public Twitter discussion instead of trying to establish these insights from a similar type of offline interaction.

Java et al., [27] and Huberman et al. [23] were among the first to study Twitter and interaction structures there. Also some other interesting cases of using conference data visualization have been carried out e.g. in IRIS conference (Information Systems Research Seminar in Scandinavia), in which the evolution of conference authors' social networks and the research topics were visualized between the time period of 1978-2006. Second, Huhtamäki et al. [24] provide an example of employing information visualization in conferences for a data-driven development of online conference workspaces, supporting conference participants before, during and after the conference. On basis of their case study, they conclude that the dynamics of the conference workspace usage can be better understood by both tracking and visualization of the usage, and thus, insight can be provided e.g. on the popularity of individual views, navigation paths, as well as, interestingly, also the structure and the development of the social network of the participants.

2.3 Visual network analysis

Recent research related to Twitter use in conferences seems to have used mainly simple quantitative analysis [15] or mapping the tweets with talks and subevents that they refer to [41]. Card et al. [11] define information visualization as "the use of computer-supported, interactive, visual representations of abstract data to amplify cognition". In recent years,

various types of information visualization approaches for social media have been created, for instance based on social network analysis methods [20,13,25], and the genre of information visualization develops rapidly.

The methods of social network analysis (SNA) provide an intuitive approach for analyzing social media [20]. Twitter users can be modeled as network nodes that are interconnected through discussions and retweets. The co-occurrence of hashtags provide another approach for taking the network approach in analyzing social media content.

Network analysis introduces a set of methods, practices and metrics for supporting the investigation and representation of social media data. Network metrics can be produced for the individual nodes of the network as well as for the network as a whole [42]. Node indegree, outdegree, betweenness and other metrics can be used to highlight nodes in different roles e.g. through node size:

Node degree value representing the number of connections a particular node has is the simplest metric for node centrality. For directed networks, node indegree, i.e. the number of connections pointing towards a node, allows the analysis of node prestige as high indegree indicated interest towards the node; outdegree indicates activity.

Node betweenness value indicates the number of shortest paths that pass a particular between any two nodes in the network; high betweenness shows that a node has a connecting role as bridge between the different parts of the overall network. Betweenness value is high for nodes having a bridging role connecting different parts of a network.

Network structure can also be analyzed. Modularity analysis, for example, allows for the identification of clusters of nodes that are more interconnected with each other than with the rest of the network [7]. Hence, clustering network nodes allows further support for insight on the internal structure of the network e.g. revealing the emergent subgroups within the network.

Freeman [17] points out the key strengths of visual network analysis: it helps investigators both in finding patterns within a networked phenomenon and in communicating the results to those that are interested.

Interactivity is a key ingredient in allowing insight through information visualization [29]. Moreover, "visual analysis typically progresses in an iterative process of view creation, exploration, and refinement." [22]

The information visualization reference model [8] defines the key steps of the technical implementation process: Raw Data is collected, refined into Data Tables, transformed into Visual Structures from which, finally, Views are created for representing the data. The Network Analysis and Visualization (NAV) process model that Hansen et al. [21] derived by observing graduate students learning social network analysis of online communities defines the key phases moves the general information visualization process towards network paradigm. After defining the analysis goals, steps of data collection and structuring, data interpretation, and report preparation complete the NAV model. The model also recognizes the need for the group of investigators to learn the SNA concepts and tools as part the analysis process.

3. Research method and approach

In this study, we apply the process of data-driven visual network analytics for providing insights on how Twitter was used before, during and after CMAD 2013. As Twitter can be seen to present an information system, we approach this study with the case study approach, which has been found to be a legitimate way of adding to the body of knowledge in the information systems field; it provides detailed and analyzed information about real world environments through examples of phenomena under research [6].

3.1. Case CMAD 2013

Our case environment is Community Manager Appreciation Day (CMAD 2013) conference. It was an event held on January 28, 2013 in Tampere, Finland as part of international conference series arranged globally every fourth Monday of January. CMAD conferences have been organized since 2010 and they originate from Jeremiah Owyang's blog [35] to recognize and celebrate the efforts of community managers around the world using social media and other tools to improve customer experiences. It was followed by a series of conferences organized at the same time in numerous cities in 2011.

The organizing committee of the second CMAD conference (CMAD 2013) in Finland included more than 100 people, with 15 people participating in the planning meetings (face-to-face or virtually by Adobe Connect). The CMAD participants can be for a significant part considered as advanced lead users of community approaches and social media. Total of 155 people participated in the CMAD 2013 [10] conference and 223 people in the online live stream [36] during the conference day.

3.2. Collection and extraction of social media data

We will concentrate here on the use of Twitter– even if also other means of social media were used in CMAD 2013 conference–for two important reasons.

First, Twitter data, unlike the data of many other commonly used social media- related application such as Facebook, is easily openly available for analysis and visualization purposes. Second, due to the restrictions and challenges of Twitter as a communication media mentioned in this study, visualization can significantly add the value of Twitter data in many respects.

In more practical and technical terms, we implemented a tailored batch script in Python (version 2.7.1) that accessed the Twitter REST API (version 1.0) to collect all the tweets sent between January 21 and March 4 that included the word cmadfi as part of their content. Twitter REST API was sufficient for collecting the tweets because it allows retrieving 1500 tweets at a time, 350 times in an hour where as more high-volume Twitter streams insist applying Twitter Streaming API instead of the REST API.

An open source NoSQL (Not Only SQL) database MongoDB (1.6.5) was used to manage and query the tweets. Whereas the volume of tweets would not insist this, we appreciate the fact that as a document database, MongoDB does not insist the specification of a schema for the collected data [8], thus allowing quick changes in processes collecting and refining the data.

For this study, we collected a total of 2686 tweets over a six-week period starting from January 21, 2013. Before the CMAD day, 162 tweets were sent, 160 with hashtag cmadfi and two in which cmadfi was found in some other form. During the CMAD day in January 28, the participants and others interested sent a total of 2138 tweets, 2126 with the hashtag cmadfi. In addition, 386 tweets were collected after January 28, 374 including the hashtag cmadfi. Figure 1 shows the cumulation of tweets over time. The blue spline shows the cumulated overall amount of tweets. The red spline represents the relative change in the amount of tweets.

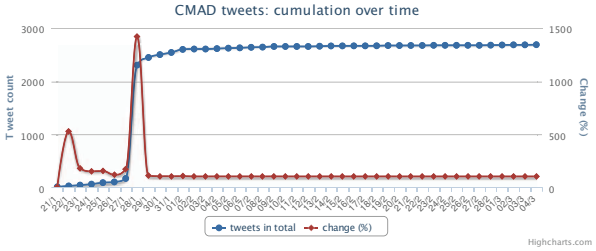

Figure 1. Cumulation of the collected tweets before, during and after CMAD 2013.

As the timeline shows, the volume of Twitter communication peaks during the conference day and remains somewhat active for the next few days. Importantly, the vast majority of the activity occurs during the conference day.

3.3. Cleaning of data in social networks

In general, Twitter data allows straightforward analysis. The used REST API puts out tweet data including both the 140 character messages, in which users (e.g. @markosuomi) and hashtags (e.g. #cmadfi) are represented with a syntax that is easy to process programmatically, and in addition includes rich metadata for each tweet, including e.g. tweet sender, time when the tweet was sent, the optional geolocation, i.e. the place where the tweet was sent etc.

A tailored Python script was implemented to identify the Twitter users that were mentioned as well as the hashtags included in each tweet. The script further transformed the refined data into three networks:

- The first network is a two-mode network including two types of nodes, representing both Twitter users and hashtags. A pair of users is connected to each other when one has mentioned the other. Users are also connected to the hashtags they have used in their tweets as well as to the hashtags that are used in the tweets they have been mentioned in.
- The second network shows the interconnections between people communicating over Twitter. More specifically, with interconnections, we refer to users mentioning each other in tweets through commenting, discussions and retweets.
- The third network represents the co-occurence of hashtags included in the tweets. The weight of the connection indicates the number of times a user has mentioned another user in a tweet.

The Python script uses NetworkX library (version 1.7) to construct the network and serialize it in Graph Exchange XML Format or GEXF (version 1.2). For temporal analysis, another Python script was implemented, to transform the data in Data Tables into timeline-based Visual Structure. Highcharts (version 2.1.9), a JavaScript-based software library for developing interactive charts, was used to implement the timelines. Python library Cheetah was used to aid the creation of the visualizations.

3.4. Data analysis and visualization

For structuring the analysis process, we applied the Network Analysis and Visualization process model. As we realize that cumulative networks alone do not allow insights on the dynamics of discussions related to CMADFI, we complemented the network analyses with temporal analysis through timeline views. For this, we followed an approach that Ebner and Reinhardt [16] introduced in which two splines are used to show the cumulation of tweets. The first spline shows the total amount of tweets over time and second

spline highlights relative change in the amount of tweets. Figure 1 shows the resulting timeline.

To facilitate the analysis, we implemented a Python script to compile a set of basic statics of the tweets, allowing all the members of the research team to conduct detailed analysis with a spreadsheet processor or some other analysis tool. This statistics include the timeframe of the sample, the number of tweets in a sample, list of the first and last tweets included in the sample, list of hashtags and their popularity, list of tweeters and the number of tweets that they have sent, and a list of the most mentioned tweeters and the number or times they are mentioned in a sample.

For analyzing and visualizing the networks, we used Gephi, an interactive visualization and exploration platform available in open source [4]. Following the NAV model, Gephi was used to layout the networks, calculate metrics for the network nodes, analyze networks for possible subnetworks or clusters and adjust the visual properties of the visualized network according to the analysis.

In this particular case, we decided to use the value of weighed node indegree to define node size. Indegree refers to the amount of connections pointing to a node, in this case the number of mentions that a particular used has received. The weighed value takes into account multiple incoming connections, i.e. connections in which a person has mentioned another are more important than individual mentions.

The layout of the networks in this study is the result of a force driven layout algorithm in which nodes repel each other and the edges connecting the nodes act as springs pulling the nodes back together [12,18]. As a result, nodes that are interconnected will be placed close to each other.

For distributing interactive versions of the network visualizations over the Web, we used Gexf.js, a Javascript-based GEXF Viewer for Gephi.

4. Results and findings

In this study, we wanted to learn more about the key users and dynamics of the interactions of the conference. We used information visualization from Twitter data to provide insights of the networks of people participating onsite and/or online in the conference for co-organizing conferences.

Many insights can be derived by looking at the network of people tweeting and their discussions topics before, during and after the conference. Figure 2 provides an overview of the people and their discussions before, during and after the conference. The interactive visual representation provides a way to observe the network through the lens of each individual person or discussion topic used in the conference. Each person can thus observe how he or she is connected to other persons and discussion topics and also see the related network metrics, such as degree, betweenness and weighted indegree.

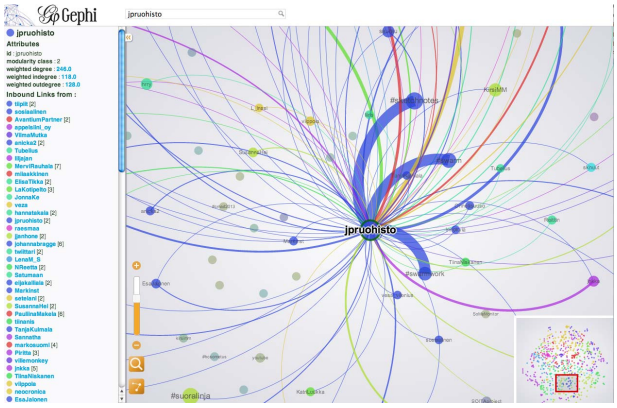

Figure 2. Snapshot of two-mode network of people tweeting and their discussion topics before, during and after the conference day. Interactive version is available: http://www.tut.fi/novi/hicss2014/

To complete the two-mode network of people and discussions topics with more specific insights, the individual one-mode networks give snapshots of the most influential people and the most popular discussion topics. Whereas the two-mode network allows a fuller overview of the discussions, one-mode networks can be used to more specifically investigate the role of different users as well as the patterns in usage of the hashtags.

The visualization of the network of people tweeting before the CMAD 2013 conference day (using #cmadfi hashtag) is illustrated in Figure 3. The network of people tweeting based on the nodes indegree uncovers the most often mentioned tweeters before the conference day. For the conference co-organizers familiar with the Twitter usernames, that were collaboratively collected and shared using Google Docs Spreadsheet, the visualization of the network of people provides many insights.

First, the network of people highlights the most influential people in the network in Twitter, the larger the influence, the larger the size of the node and node label (Twitter username). Second, the interests of the people are made visible by the connections to other people, the larger the interest the larger the size of the connection (line width in Figures 3-5). The node color represents the clusters of nodes in the network, as according to a community-detection algorithm that analyzes the network to find groups of nodes that are particularly tightly interconnected.

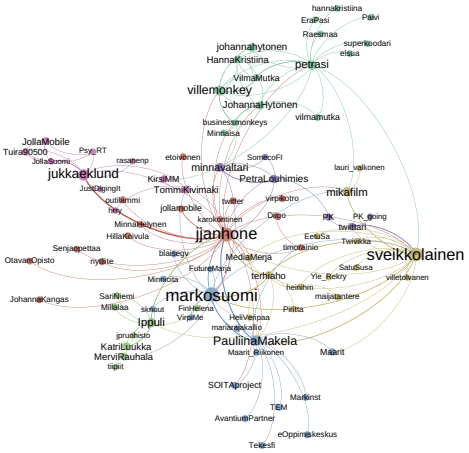

Figure 3. Force driven network of people tweeting before the conference day.

In the network of people before the conference in Figure 3, the most influential people in Twitter include co-organizers (@jjanhone, @PauliinaMakela), speakers (@markosuomi, @jukkaeklund, @mikafilm), and potential participants (@villemonkey, @petrasi, @ippuli, @minnavaltari, and @PetraLouhimies). The visualization helps to uncover clusters in the network (Figure 3).

By looking at the messages exchanged in the cluster it is possible to determine how the clusters are formed. One cluster is formed around @petrasi, @villemonkey, @JohannaHytonen and @VilmaMutka that discussed for example who they know that are going to attend the conference and practical arrangements such as transportation from Jyväskylä to Tampere for the conference. Similarly local clusters formed around people who discuss about going to the conference from Helsinki people going to the conference from Turku. Two clusters form around influential co-organizers @jjanhone and @Pauliina that for instance inform about the coming conference and the status of the conference arrangements and provide information about community managers. In addition clusters form around influential presenters (@markosuomi and @jukkaeklund).

In the network in Figure 4, a few clusters emerge. The top 10 most influential people in Twitter based on node size during the conference day include speakers (e.g. @toninummela, @markosuomi, and @senjalarsen), participants (e.g. @sveikkolainen), co-organizers that were also speakers (@PauliinaMakela and @jjanhone) and a representative (@MerviRauhala) from a sponsor organization (@Markinst).

By knowing which people were speakers in the conference (e.g. @toninummela, @markosuomi, @senjalarsen, @AnttiIsokangas, @jpruohisto, and @mikafilm), it is possible to deduce, at least in terms of Twitter presence, which presentations raised most discussion and were most interesting to the audience.

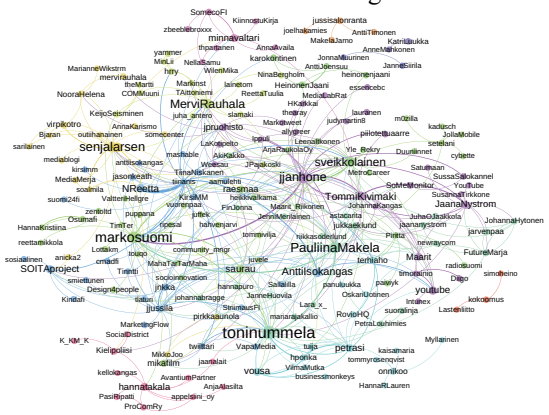

Figure 4. Force driven network of people tweeting during the conference day.

The most important observation, however, is the tightly interconnected nature of the overall network.

Comparing Figures 3 and 5 reveal that discussions were more active and more people were involved in the discussions after the conference than before the conference.

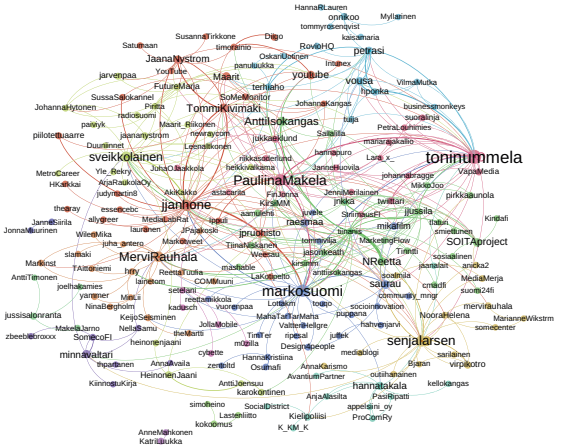

Figure 5. Force driven network of people tweeting after the conference day.

The discussions about conference presentations continued after the conference and a relatively large group of people participated in the discussions after the conference, considering that only 155 people attended the conference onsite.

The visualization of the network of hashtag co-occurence before the conference day uncovers, which discussion topics were associated with the conference and the popularity of those discussion topics before the conference (Figure 6).

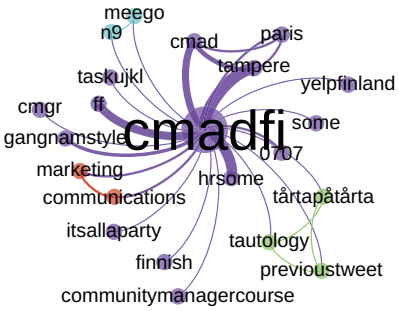

Figure 6. Force driven network of hashtags before the conference day.

The most popular discussion topics before the conference day (Figure 6) related to, Community Manager Appreciation Day (#cmad) conferences in Tampere and Paris, different involved communities, such as #hrsome and #meego, follow friday (#ff) that is suggesting interesting Twitter users and helping them get followers, and community managers in general (#cmgr, #communitymanagercourse).

The network of hashtags during the conference day is illustrated in Figure 7.

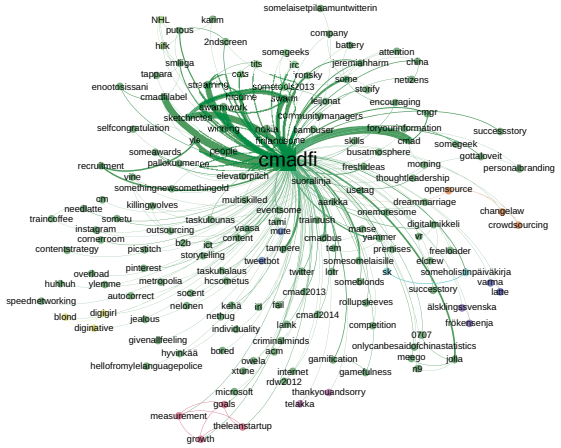

Figure 7. Force driven network of hashtags during the conference day.

During the conference the third most discussed topic after #cmadfi and #cmad was #cmadfilabel, chosen and voiced during conference presentation [26] for discussing the appropriate translation of "community manager" in Finnish. Most of the top 10 hashtags used referred to the presentations, such as #sketchnotes that included notes in the form of visual drawings of the presentations, or more directly to the concepts used in the presentations, such as #streaming, #swarm, or the context of the presentation (e.g. #yle).

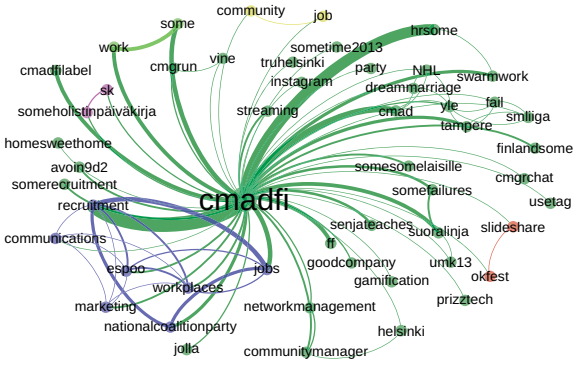

Figure 8. Force driven network of hashtags after the conference day.

The hashtags included in the discussions after the conference are shown in Figure 8. The most popular discussions topics include recruitment of community managers (#recruitment), #hrsome community, Community Manager Appreciation Day (global) conferences (#cmad), and jobs for community managers (#jobs, #job, #work).

5. Discussion

Current literature brings forth the possibilities and benefits but also the restrictions of Twitter as an approach to support conference organization, conference-related collaboration and conference content sharing and communication. We wanted to understand and add to that discussion with a case study of CMAD 2013 and visualizing its Twitter activity.

Overall, network visualizations were useful in revealing the overall structure of the communications that occur during conferences.

For Twitter data during conference, the visualization of the network of people uncovered the influential people involved in the conference as co-organizers, speakers, sponsors or participants. The most influential speakers indicate the most interesting conference presentations and discussion themes for the conference audience, uncover various conference topic- related needs of the conference attendees, and provide insight into planning the contents and speakers of the next conference.

The visualization of Twitter hashtag networks provides several insights of the discussion themes during the conference. One of the findings was that the discussions tend to scatter when hashtags are created bottom-up, each person creating his or her own hashtags to describe content (e.g concepts) of the presentations. While those hashtags that were created and communicated top-down either by conference co-organizers or speakers were more popular and seemed

to last longer. To confirm that the discussions using top-down created hashtags lasted longer, would however need further analysis [cf. 10].

An important finding from the visualization of hashtag networks during conference was that from the collected Twitter data the content links that were created bottom-up, e.g. by the conference participants, in some cases led to broken links or discontinued services and thus missing conference related content.

Information visualization of Twitter data before the conference had interesting implications for conference co-organization. For instance, the clustering of the network of people pointed out to similar groups or networks of people. The content of tweets revealed that some of the similar groups are actually local clusters or networks from different cities that are planning to attend the conference, and discuss transportation related issues. The before the conference discussions indicate that conference co-organizers should take steps in improving the information about transportation to the conference and consider instructing, facilitating or arranging transportation related issues.

The identification of the influential people in Twitter have several implications for conference co-organization. Due to their networks and prestige, the discovered most influential participants can also be considered as potential co-organizers and co-marketers for the future conferences. Furthermore, the discovered connections between the attendees can be used by the co-organizers to identify people with similar interests and for example plan sessions than interest certain groups of people.

The clustering of nodes in the people network can help in observing emerging discussion groups sharing a similar novel interest in a particular topic. Reasons for the emergence of such groups include e.g. spin-off discussions that a particular presentation catalyzed but did not fully cover (a possible topic for future conference rounds) or problems in accessing online content with a particular device or software. At best, such groups could be investigated in detail with the help of supporting views showing the topics or the actual contents of the particular discussion.

Many of the discussions after the conference connected the CMAD 2013 Twitter stream to many other existing streams, thus adding its visibility. In total, three times the amount of the participants were involved in the discussions. The reported media coverage allows organizers to better observe the impact of the conference.

This study leaves room for future studies in several areas. First, visualizations cover a huge amount of analysis techniques, and only a fraction has been used and investigates in this study. For instance, to gain further insight, the following visual network analyses would be highly useful: 1) analysis of number of tweets per user and average posts per user [15] to understand what percentage of users are actually active contributors and do the contributions follow e.g. the 90-9-1 rule [34], 2) the growth of the network of followers of conference participants, e.g. compared before and after conference to better understand the impact of the conference for participants 3) analysis of preferential attachment [2,5] for identifying the nodes with strategic position in the network and for investigating to what degree is the growth of network of people and network of discussion due to higher probability of people that have large number of connections or due to their activity in the conference.

Second, further avenues of research could include other types of available social media data, such as Facebook data of conference organization, should be combined in the analyses to provide a more varied and detailed picture of the prestige of persons, presentations and discussions in conferences, because quite naturally, not all significant discussions take place in Twitter, and some influential persons may be active elsewhere than in an individual communication channel such as Twitter.

References

[1] Aarts, O., van Maanen, P.-P., Ouboter, T., and Schraagen, J.M. Online Social Behavior in Twitter: A Literature Review. *Data Mining Workshops (ICDMW), 2012 IEEE 12th International Conference*, (2012), 739–746.

[2] Abbasi, A. and Hossain, L. Investigating Preferential Attachment Behavior over the Evolution of Disaster Response Networks. *46th Hawaii International Conference on Systems Science (HICSS-46)*, (2013).

[3] Ahuja, G., Soda, G., and Zaheer, A. The genesis and dynamics of organizational networks. *Organization Science 23*, 2 (2012), 434–448.

[4] Amiando. *Social Media & Events Report 2012. How is the event industry using Social Networks?* 2012.

[5] Barabási, A.-L. and Albert, R. Emergence of scaling in random networks. *science 286*, 5439 (1999), 509–512.

[6] Benbasat, I., Goldstein, D.K., and Mead, M. The case research strategy in studies of information systems. *MIS quarterly*, (1987), 369–386.

[7] Blondel, V.D., Guillaume, J.-L., Lambiotte, R., and Lefebvre, E. Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment 2008*, 10 (2008), P10008.

[8] Bonnet, L., Laurent, A., Sala, M., Laurent, B., and Sicard, N. Reduce, you say: What nosql can do for data aggregation and bi in large repositories. *Database and Expert Systems Applications (DEXA), 2011 22nd International Workshop on*, (2011), 483–488.

[9] Boyd, D., Golder, S., and Lotan, G. Tweet, tweet, retweet: Conversational aspects of retweeting on twitter.

System Sciences (HICSS), 2010 43rd Hawaii International Conference on, (2010), 1–10.
[10] Bruns, A. How long is a tweet? Mapping dynamic conversation networks on Twitter using Gawk and Gephi. *Information, Communication & Society 15*, 9 (2012), 1323–1351.
[11] Card, S.K., Mackinlay, J.D., and Shneiderman, B. *Readings in information visualization: using vision to think*. Morgan Kaufmann Pub, 1999.
[12] Chau, M. and Xu, J. Mining communities and their relationships in blogs: A study of online hate groups. *International Journal of Human-Computer Studies 65*, 1 (2007), 57–70.
[13] Chen, B. Is the Backchannel Enabled? Using Twitter at Academic Conferences. (2011).
[14] DeVoe, K.M. "You'll Never Guess Who I Talked To!": Tweeting at Conferences. (2010).
[15] Ebner, M., Mühlburger, H., Schaffert, S., Schiefner, M., Reinhardt, W., and Wheeler, S. Getting granular on twitter: tweets from a conference and their limited usefulness for non-participants. In *Key Competencies in the Knowledge Society*. Springer, 2010, 102–113.
[16] Ebner, M. and Reinhardt, W. Social networking in scientific conferences–Twitter as tool for strengthen a scientific community. *Proceedings of the 1st International Workshop on Science*, (2009).
[17] Freeman, L.C. Methods of social network visualization. In R.A. Meyers, ed., *Encyclopedia of Complexity and Systems Science*. Springer, Berlin, 2009.
[18] Fruchterman, T.M. and Reingold, E.M. Graph drawing by force-directed placement. *Software: Practice and experience 21*, 11 (1991), 1129–1164.
[19] Grimnes, G. Semantic Conference Program. (2006).
[20] Hansen, D., Shneiderman, B., and Smith, M.A. *Analyzing social media networks with NodeXL: Insights from a connected world*. Morgan Kaufmann, 2010.
[21] Heer, J. and Boyd, D. Vizster: Visualizing online social networks. *Information Visualization, 2005. INFOVIS 2005. IEEE Symposium on*, (2005), 32–39.
[22] Heer, J. and Shneiderman, B. Interactive dynamics for visual analysis. *Queue 10*, 2 (2012), 30.
[23] Huberman, B., Romero, D., and Wu, F. Social networks that matter: Twitter under the microscope. *Available at SSRN 1313405*, (2008).
[24] Huhtamäki, J., Nykänen, O., and Salonen, J. Catalysing the development of a conference workspace. In *Online Communities and Social Computing*. Springer, 2009, 43–52.
[25] Huhtamäki, J. and Parviainen, O. Verkostoanalyysi sosiaalisen median tutkimuksessa. In S.-M. Laaksonen, J. Matikainen and M. Tikka, eds., *Otteita verkosta - Verkon ja sosiaalisen median tutkimusmenetelmät*. Vastapaino, Tampere, 2013.
[26] Janhonen, J. and Mäkelä, P. CMAD Suomessa ja maailmalla – keitä ne on ne community managerit ja mitä he oikein tekevät? 2013. http://www.videonet.fi/hermia/20130128/11/#.
[27] Java, A., Song, X., Finin, T., and Tseng, B. Why we twitter: understanding microblogging usage and communities. *Proceedings of the 9th WebKDD and 1st SNA-KDD 2007 workshop on Web mining and social network analysis*, (2007), 56–65.
[28] Jussila, J., Huhtamäki, J., Kärkkäinen, H., and Still, K. Information visualization of Twitter data for co-organizing conferences. *AcademicMindTrek '13: International Conference on Making Sense of Converging Media Proceedings*, (2013).
[29] Kosara, R., Hauser, H., and Gresh, D.L. An interaction view on information visualization. *State-of-the-Art Report. Proceedings of EUROGRAPHICS*, (2003).
[30] Lerman, K. and Ghosh, R. Information Contagion: An Empirical Study of the Spread of News on Digg and Twitter Social Networks. *ICWSM 10*, (2010), 90–97.
[31] Letierce, J., Passant, A., Breslin, J., and Decker, S. Understanding how Twitter is used to spread scientific messages. (2010).
[32] McKendrick, D.R., Cumming, G.P., and Lee, A.J. Increased Use of Twitter at a Medical Conference: A Report and a Review of the Educational Opportunities. *Journal of medical Internet research 14*, 6 (2012), e176.
[33] McNeill, T. More than just passing notes in class? Reflections on theTwitter-enabled backchannel. http://www.scribd.com/doc/16287533/More-than-just-passing-notes-in-class-The-Twitterenabled-backchannel#download.
[34] Nielsen, J. Participation inequality: Encouraging more users to contribute. *Jakob Nielsen's alertbox*, 2006. http://www.nngroup.com/articles/participation-inequality/.
[35] Owyang, J. Community Manager Appreciation Day #CMAD (Every 4th Monday of Jan) | Web Strategy by Jeremiah Owyang | Social Media, Web Marketing. 2010. http://www.web-strategist.com/blog/2010/01/25/community-manager-appreciation-day-cmad-every-4th-monday-of-jan/.
[36] Rainio, T. Hermia webcast. 2013. http://www.videonet.fi/hermia/20130128/.
[37] Reinhardt, W., Ebner, M., Beham, G., and Costa, C. How people are using Twitter during conferences. *Proceedings of 5. Edumedia Conference*, (2009), 145–156.
[38] Reinhardt, W., Maicher, J., Drachsler, H., and Sloep, P.B. Awareness support in scientific event management with ginkgo. *Proceedings of the 11th International Conference on Knowledge Management and Knowledge Technologies*, (2011), 40.
[39] Ross, C., Terras, M., Warwick, C., and Welsh, A. Enabled backchannel: Conference Twitter use by digital humanists. *Journal of Documentation 67*, 2 (2011), 214–237.
[40] Sopan, A., Rey, P.J., Butler, B., and Shneiderman, B. Monitoring Academic Conferences: Real-time Visualization and Retrospective Analysis of Backchannel Conversations. *Social Informatics (SocialInformatics), 2012 International Conference on*, (2012), 62–69.
[41] Stankovic, M., Rowe, M., and Laublet, P. Mapping tweets to conference talks: A goldmine for semantics. *Workshop on Social Data on the Web, Shanghai, China*, (2010).
[42] Wasserman, S. and Faust, K. *Social network analysis: Methods and applications*. Cambridge university press, 1994.
[43] Williams, S.A., Terras, M., and Warwick, C. What people study when they study Twitter: Classifying Twitter related academic papers. *Journal of Documentation 69*, 3 (2013), 1–74.