School of Mathematical
and Physical Sciences

University of Reading

Department of Mathematics and Statistics

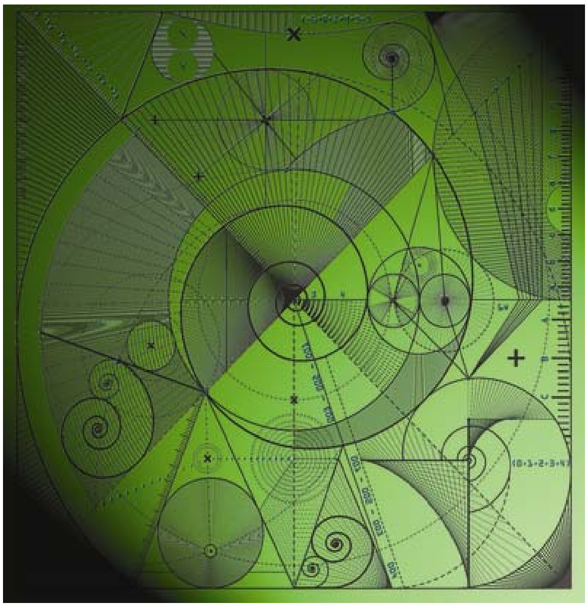
Preprint MPS-2012-25

14 November 2012

Structure, pace and balance in Twitter conversations

by

Danica Vukadinovic Greetham and
Jonathan A. Ward



Structure, pace and balance in Twitter conversations

Danica Vukadinović Greetham, Jonathan A. Ward
Centre for the Mathematics of Human Behaviour
Department of Mathematics and Statistics
University of Reading, Whiteknights, RG6 6AX, UK

Abstract

Twitter is both a micro-blogging service and a platform for public conversation. Direct conversation is facilitated in Twitter through the use of @'s (mentions) and replies. While the conversational element of Twitter is of particular interest to the marketing sector, relatively few data-mining studies have focused on this area. We analyse conversations associated with reciprocated mentions that take place in a data-set consisting of approximately 4 million tweets collected over a period of 28 days that contain at least one mention. We ignore tweet content and instead use the mention network structure and its dynamical properties to identify and characterise Twitter conversations between pairs of users and within larger groups. We consider conversational balance, meaning the fraction of content contributed by each party. The goal of this work is to draw out some of the mechanisms driving conversation in Twitter, with the potential aim of developing conversational models.

1 Introduction

The rapid uptake of online social media, combined with consumer behavioural changes around television and news broadcasting, has instigated a sea change in attitudes within the advertising and marketing sectors. A frequently encountered adage is that "everything is about conversation and not about broadcasting" [9, 5]. By facilitating public addressability through the @ sign (so called 'mentions') and enabling private messages, Twitter has confirmed their intention to function as a communication channel as well as a broadcasting tool. Access to large quantities of data produced by Twitter users has resulted in a surge of interest from the academic community, who have largely focused on Twitter's information flow and retweet behaviour, and hence implicitly the underlying network of 'followers' (e.g. [18, 17]). While broadcasting short messages, or micro-blogging, remains an important component of Twitter use, to our knowledge comparatively little work has addressed the mining of (public) conversations on a large scale. Consequently, we focus in this paper on analysing the network of communication patterns resulting from mentions in Twitter.

Inspired by the importance of pauses between exchanges in face-to-face conversations, drama and music, our main aim has been to explore the structure and rhythm of typical Twitter conversations. Although it may not always be clear, even from message content, what intention a user had in mind when posting—information seeking or information sharing, broadcasting or conversation—we have tried to specifically extract conversations by focusing our data-analysis on reciprocated tweets. Moreover, we completely ignore the content of conversations and concentrate on structural and dynamic properties of the underlying mentions network. Our main objective was to mine actionable insights that could inform our knowledge of conversational mechanisms and the frequency/timings of tweets. Our hope is that quantifiable insights from this analysis could inform a simple, data driven model of the timing and structure of Twitter conversations.

A large number of registered Twitter accounts are operated by automated software scripts, known as *bots* [16]. While such accounts are encouraged for the purpose of developing applications and services, bots whose functions violate Twitter policy (e.g. spammers) are common. The analysis of conversational patterns and the development of associated models have potential application for those trying to develop algorithms that can identify nuisance bots. Furthermore, the identification of groups of Twitter users who, through conversational behaviour, are particularly influential on a specific topic would be particularly attractive in the marketing sector. Thus, understanding conversational structure could impact the design and implementation of social media campaigns and potentially provide a quantitative comparison between Twitter discourse and other channels of communication, such as face-to-face, telephone, SMS, forums or email. We hope that studying Twitter conversation can ultimately improve user experience.

In Section 2, we give an account of previous work in this space. The Twitter data-set and its general features are described in Section 3. Our results of pairwise and multiple conversations are presented in Section 4. Finally, in Section 5 we summarise and describe possible directions of future work.

2 Previous work

The phenomenal uptake of Twitter over the last few years has resulted in a rapidly growing interest in mining Twitter data and particularly sentiment analysis of tweets. A recent study analyzing a large amount of Twitter and Facebook data [11] found correlations between friendship/follower relations and positive/negative moods of Twitter users. Diurnal and seasonal mood rhythms that are common across different cultures have also been identified in cross-cultural Twitter data [4], shedding light on the dynamics of positive and negative affect. Connections between emotions inferred from tweets and the Dow Jones index made headlines [1], as an interesting and potentially useful observation of society.

A study of conversations within a sample of 8.5k tweets collected over an hour long period [8] found that the @ sign appeared in about 30% of the collected sample, its function was mostly for addressing (as intended) and it was relatively well reciprocated—around 30% of messages containing an @ were reciprocated within an hour. The majority of these conversations were short, coherent exchanges between two people, but longer exchanges did occur, sometimes consisting of up to 10 people. They found that

> *"…Tweets with @ signs are more focused on an addressee, more likely to provide information for others, and more likely to exhort others to do something—in short, their content is more interactive. In contrast, tweets without @ signs are more self-focused, although they also report other's experiences, and they make more general announcements."*

Although our collecting method was completely different (we focused on tweets containing an @ sign exclusively over a much longer interval), it is instructive to compare our results with the findings described above.

In other work concerning Twitter conversation [12], a relatively large corpus and content (topic) analysis of 1.3million tweets was used to develop an unsupervised model of dialogue from open-topic data. In our work we completely ignore content, enabling us to focus on timing, the structure and balance of conversation (in particular the order of tweets between pairs of individuals) and multi-users conversations which could help to improve statistical models of dialogue.

3 Analysis

3.1 Data

The Twitter data-set investigated in this paper was collected on our behalf by Datasift, a certified Twitter partner, allowing us to access the full Twitter firehose rather than being rate-limited by the API. The data-set consists of all UK based[1] Twitter users that sent tweets with at least one mention between 8 Dec 2011 and 4 Jan 2012 (28 days in total). In the remainder of the paper, use of the word 'tweet' will specifically mean tweets containing at least one mention. Mentions are messages that include an @ followed by a username. Thus if person a puts "@b", it designates that a is addressing the tweet to b specifically. Mentions are not private messages and can be read by anyone who searches for them. A tweet can be addressed to several users simultaneously using @ repetitively. We preprocessed the data, removing empty mentions and self-addressing[2] and created a directed multigraph, or mentions network, containing $3,614,705$ timestamped arcs (individual mentions) from a total of $819,081$ distinct usernames, or nodes. Of these distinct usernames, $732,043$ were "receivers", i.e. to whom a message was addressed, and $137,184$ were "tweeters", i.e. people who tweeted a message with a mention. There were approximately 50k nodes that appeared both as tweeters and receivers. Note that our graph is a multigraph, meaning that multiple arcs are allowed between pairs of nodes, each having a direction and timestamp.

[1] All Twitter users appearing in our data-set had selected the UK as their location.

[2] Self-mentioning was surprisingly common in the data-set: 12,680 different users created a total of 44,319 self-mentions, with the maximum being 5,586 from an automated service that advertises itself at the end of each tweet.

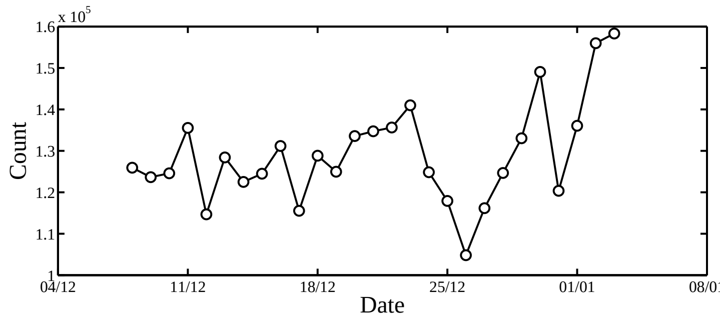


Figure 1: Time series of total daily tweet counts.

3.2 Aggregate daily tweet behaviour

We start by giving an overview of the daily tweet behaviour captured in our data-set. We focus here on tweets, which may contain one or more mention. The average total number of tweets sent by all users per day was $129,096$ with a standard deviation of $11,945$. A time series of the total number of tweets sent each day is plotted in Fig. 1. The 26 Dec 2011 (Boxing Day) has the lowest total tweet count at 104,808 and the last day, 4 Jan 2012, has the largest total tweet count at 158,319.

The largest number of tweets sent by a single user on a given day was 590; the largest number of tweets sent by a single user over the 28 days was 5,604; the mean daily tweet rate (from usernames that tweeted) was 0.9410. In Fig. 2 we plot the distribution of mean daily tweet counts per user and compare with a fitted exponential distribution (grey line). This indicates that the distribution has a 'heavy tail' (but not power-law), meaning that while most people tweet less than once a day, some people have extremely high average daily tweet counts. We observe that a high tweet rate is not necessarily indicative that an account is being operated by a bot, making the task of identifying such automated users non-trivial.

3.3 Time intervals between tweets

We plot the distribution of time intervals between users' tweets dt in Fig. 3(a) and (b). The largest time intervals are much less than the total number of hours covered by the 28 days (40,320) in which the data was recorded, meaning that people who tweet do so regularly. Note also the consistent peaks around time periods that are multiples of 24hrs. This indicates that

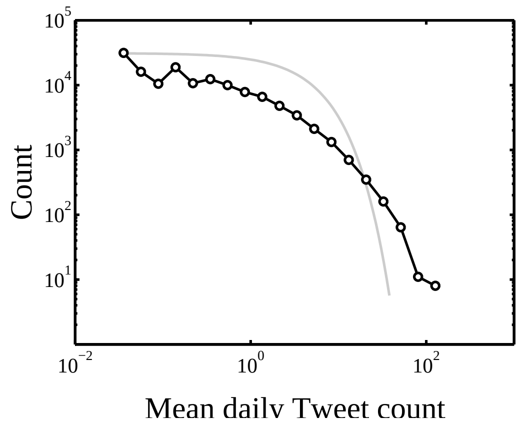


Figure 2: Distribution of mean daily tweet counts per user.

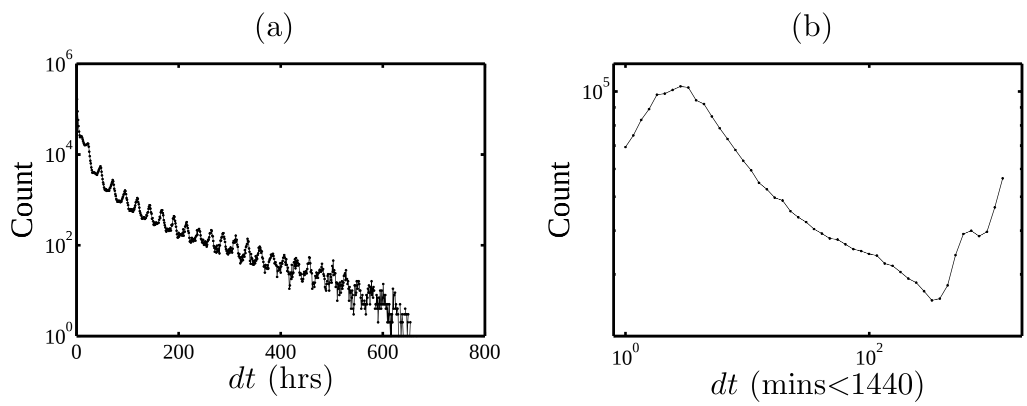


Figure 3: Distribution of time intervals between consecutive tweets from individual users

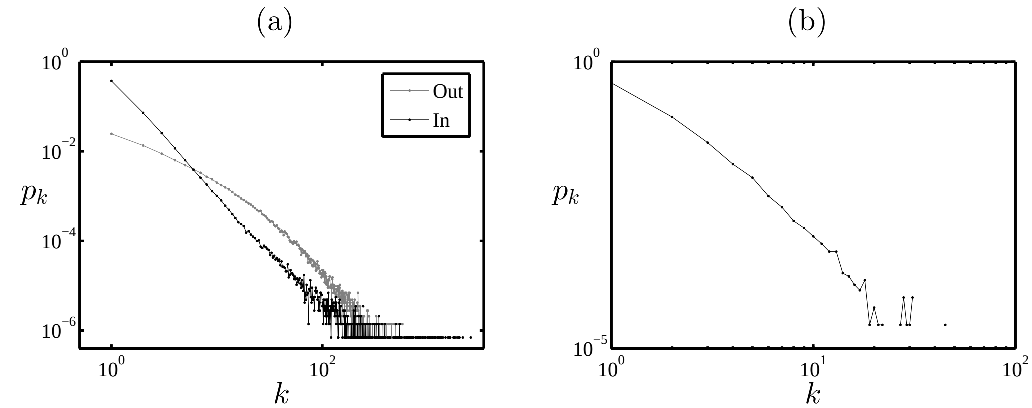


Figure 4: Panel (a): In- and Out-degree distributions in the directed adjacency matrix. Panel (b): Degree distribution in the symmetric part of the directed adjacency matrix.

people have regular times of day that they typically spend using Twitter. In Fig. 3(b) we plot the distribution of time intervals that are less than 24hrs. The mode is at about 3 minutes, which is in accord with [8], where the average and median time between exchanges were found to be 6 minutes 43 seconds and 4 minutes 24 seconds respectively. There is also a secondary peak at about 10hrs, which may reflect tweet behaviour either side of daily sleep patterns.

3.4 Mentions network

We now consider the multigraph mention network. In the first instance, we aggregate the data over the entire data collection period in order to identify structural features in what we infer as the static underlying friendship network. This is a directed weighted graph, where edge weights correspond to the total number of mentions between pairs of nodes. The maximum number of mentions sent by a single user (i.e the maximum out degree weight) is 1,171 and the maximum total number of exchanges between two users is 1,415. These are not automated accounts. The pattern of user acquaintances can be represented by a binary directed adjacency matrix A, in which an element A_{ij} is equal to 1 if i mentions j at least once in the data-set, and zero otherwise. In Fig. 4(a), we plot the in- (black) and out- (grey) degree distributions of A. These distributions are qualitatively different, the

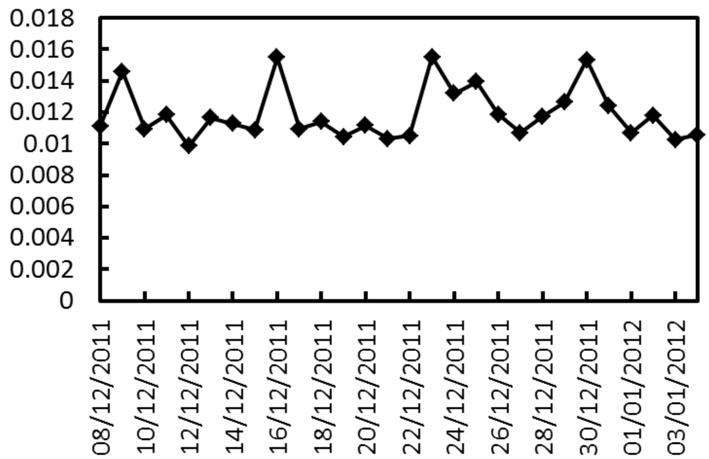


Figure 5: Clustering coefficients for daily graphs

in-degree appearing to be scale-free, but the out-degree being sub-power law. This can be attributed to the fact that individuals choose how many people to mention, but not how many people mention them. In Fig. 4(b) we plot the degree distribution of reciprocated mentions (i.e. the symmetric part of the adjacency matrix). This illustrates that users are not typically following a large number of people while simultaneously being followed by a large number of people.

To examine temporal effects in the data, we now consider daily aggregations of the full multigraph mention network, giving a series of 28 directed network snapshots. On each of these daily sub-networks, we calculated the average clustering coefficient (Fig. 5) in order to compare them over different days.

Interestingly, there are regular peaks in the clustering coefficient on Fridays. It is not clear why the mention network may be more clustered on Fridays, but one hypothesis is that people may be more inclined to broadcast jokes or interesting facts at the end of the working week to their colleagues and friends, which then triggers higher retweet rates and conversations between friends' of friends.

4 Results

We now present the main results of our analysis of pairwise and group conversations.

4.1 Conversations

An important feature of both face-to-face conversation [14, 13] and computer-mediated communication [7], is the process of turn-taking. Thus in sequences of mentions between pairs of users, say a and b, we might expect that sequences like $ABABAB$ would be more common than say $AAABBB$, where we use A to denote that party a mentions party b and likewise B to denote that party b mentions party a.

To establish if this is the case, we assume the null hypotheses that contributions are independent events with probability P_A that party a contributes to a conversation and thus probability $P_B = 1 - P_A$ that party b contributes. For a given interaction sequence of length N between parties a and b, we are interested in the number of occurrences of B following A and vice-versa. We call these *transitions*, thus the sequence $ABAABBA$ of length $N = 7$, has 4 transitions. Note that we focus on reciprocated interactions, meaning that each party makes at least one contribution and consequently that there is by default at least one transition in all interactions that we consider. We call the remaining transitions the *excess transitions*. For any sequence of length N, the maximum possible number of excess transitions is clearly $N - 2$. Under the null hypotheses, excess transitions occur with probability $P_T = 2P_A(1 - P_A)$. Since we assume that transitions are independent, the probability distribution of a given number of excess transitions is binomial, and thus the expected number is $E_T = (N - 2)P_T$ with variance $V_T = (N - 2)P_T(1 - P_T)$.

To test the null hypothesis, we consider all reciprocated pairwise interaction sequences in our Twitter data-set. For each sequence having n_X contributions from party $X \in \{A, B\}$, we assume that the probability of party a contributing is simply $n_A/(n_A + n_B)$. This does not yield any problematic probabilities (i.e. 0 or 1) since both parties always make at least one contribution.

Each sequence may have a different number of interactions and a different transition probability, but assuming that the pairwise interactions are independent, the expectation and variance of the ensemble is simply equal to the

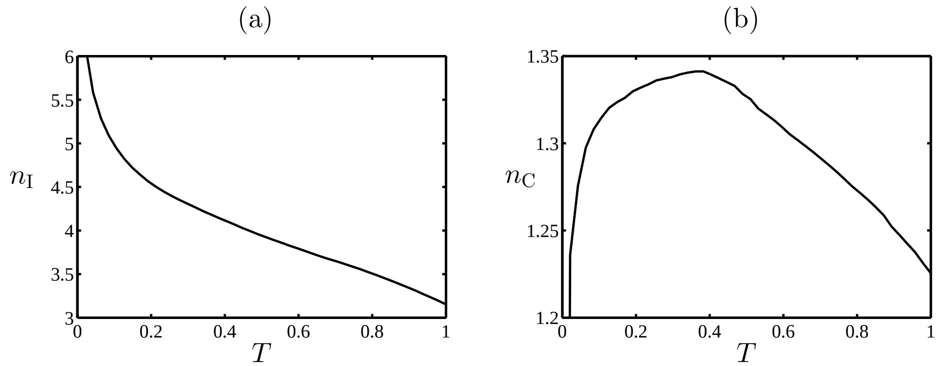


Figure 6: Panel (a): Mean number of subsequences for a range of threshold values. Panel(b): Mean number of distinct conversations for a range of threshold values

sum of the interaction expectations and variances respectively. Doing this, we find that the expected number of transitions is 85,390 with a standard deviation of 226.3, but we observe 88,758 transitions in practice, more than 15 standard deviations above the expected value. We take this as strong evidence that we can reject the null hypothesis and thus infer that the data contains a significant level of turn-taking and hence conversation.

Each sequence of pairwise interactions may constitute a number of different conversations, but ascertaining when one conversation ends and another begins may be an extremely difficult task, especially when the goal is to apply an automated processes to a large data-set. Instead of using a time-intensive lexical analysis, we investigate whether we can detect conversations by applying a simple threshold rule to the time gap between responses, where we assume that a time gap that is larger than the threshold indicates the start of a new conversation.

This method requires that we can identify a suitable threshold. To achieve this, we divide each sequence of pairwise interactions up according to a given threshold, then define distinct conversations to be reciprocated sub-sequences, i.e. sequences containing a contribution from both parties. Thus the number of sub-sequences n_{I} is always larger than the number of distinct conversations n_{C}. In Fig. 6(a) and (b) we plot the mean number of sub-sequences and the mean number of distinct conversations respectively over a range of threshold values. The number of distinct conversations n_{C} has a peak value at approximately 9hrs. This peak is expected, since we only

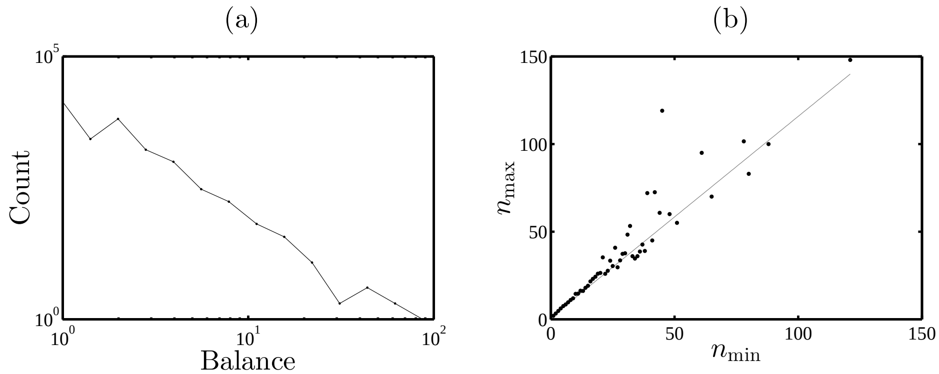


Figure 7: Panel (a): Distribution of conversation balance. Panel (b): Mean maximum conversation contribution as a function of minimum contribution.

count reciprocated interactions as distinct conversations. Thus small threshold values, which split an interaction sequence up into a large number of short sub-sequences (see Fig. 6(a)), result in relatively few distinct conversations because many of the sub-sequences feature contributions from only one party. High threshold values also result in a small number of conversations, but this is simply because they do not split the sequence up into many sub-sequences. Thus the maximum at 9hrs is a natural choice of threshold and corresponds to one's intuition that conversations may reflect diurnal patterns.

We now consider whether the number of contributions from each party are similar, or 'balanced' within pairwise interactions and conversations. For a given interaction sequence, there are two ways to compute balance, we can either consider the ratio of means $b_{\mathrm{I}} = \langle \max(n_A, n_B) \rangle / \langle \min(n_A, n_B) \rangle$ or the mean of ratios $\beta_{\mathrm{I}} = \langle \max(n_A, n_B) / \min(n_A, n_B) \rangle$. Since we only consider reciprocated interactions, both quantities are well-defined and we would generally expect $\beta > b$. For the total number of interactions between pairs, we find that $b_{\mathrm{I}} = 2.424$ and $\beta_{\mathrm{I}} = 3.457$. Thus on average, one party contributes around 3 times as much as the other. For conversations, we find that $b_{\mathrm{C}} = 1.148$ and $\beta_{\mathrm{C}} = 1.425$. These are much closer to 1, and hence more what we would expect from typical, balanced conversations. The distribution of conversation contribution ratios is plotted in Fig. 7(a), which illustrates that conversations are most likely to be balanced, but some extremely unbalanced conversations do occur. In Fig. 7(b), for each minimum conversation contribution $n_{\min} = 1, 2, 3, \ldots$, we compute the mean of the maximum contribution $n_{\max}$. This illustrates that there is a roughly linear trend and the grey line

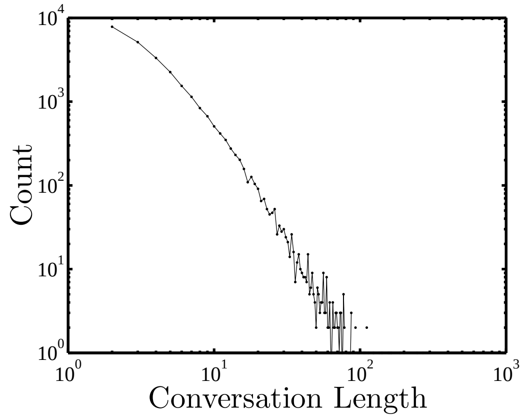


Figure 8: Distribution of conversation length.

is $n_{\max} = 1.148 n_{\min} + 1$.

The mean and median number of tweets in a conversation are 13.09 and 4 respectively, but the distribution is heavy tailed (see Fig. 8).

4.2 Multi-user conversations

By allowing multiple @ signs in one message, a Twitter user could send a tweet to several recipients simultaneously, facilitating multi-user conversations or *multicasting*. Note that because of the 140 character limit there is a physical limit on how many users each message can be multicast to.

In this part of analysis, our aim is to

- Identify multi-users exchanges;
- Find out how often they occur in our sample;
- Determine how many users typically engage in them;
- Identify their time-frame and pace;
- Calculate how balanced they are.

In order to identify multi-user exchanges using the representation of the whole data-set as a directed multi-graph G, we firstly ran non-recursive version of the Tarjan's algorithm [15, 10] as implemented in NetworkX [6] that gave us a list of the strongly-connected components of G. A directed graph is called strongly-connected if there is a path from each vertex in the graph to every other vertex. This means that for two vertices a and b there is a path in both directions, i.e. from a to b and also from b to a. Strongly-connected components of a graph are maximal subgraphs that are strongly-connected.

Pairwise conversations were discussed in sub-section 4.1, so we excluded all strongly-connected components of size 2 from the present analysis. Each strongly-connected component was then transformed into an undirected multi-graph and we ran the NetworkX implementation of the modified Bron's algorithm [2] for finding all maximal cliques in an undirected graph. Maximal cliques are the largest complete subgraphs containing a given node. We then disregarded all cliques of size two. We found in total 5,569 cliques of size 3, 4, 5, 6 and 7. The number of instances of each clique-size found within G are illustrated in Fig. 9. Clearly clique sizes larger than 4 are extremely rare. The total number of users in these cliques was 6,963 which is around 5.1% of users that tweeted. Most users were involved in just one clique but some were involved in multiple cliques. The users' involvement in cliques is illustrated in Fig 10.

When examining the time-frame of multi-user exchanges, we found that the total number of exchanges between clique members was inversely proportional to the average difference between consecutive exchanges (see Fig 11). This was not surprising, since we would expect lively conversations (with lots of exchanged messages) to have a relatively fast pace, in contrast to casual exchanges in which the differences between messages would be longer. We also found that multi-user exchanges happened over the whole 28 days period. Very few exchanges were separated by periods longer than a day, most exchanges being relatively fast paced. In Fig. 12 we plot the sorted medians between two consecutive messages within cliques. The median and mean of the medians were respectively 2,301 and 32,178 seconds. Keeping in mind that we are looking at exchanges over the whole period of 28 days (thus, theoretically an exchange could contain only six messages, have one messages on the day 1 and then one message each five days) the fact that the median is less than 1hr and the mean less than 9hrs both being much less than we might have expected, confirms that replies happen relatively quickly.

We also investigated how balanced multi-user exchanges were, although this situation is more complicated than in the pairwise case.

Firstly, we looked at the difference between the number of tweets received and sent by individual clique members. For each node, we computed the dif-

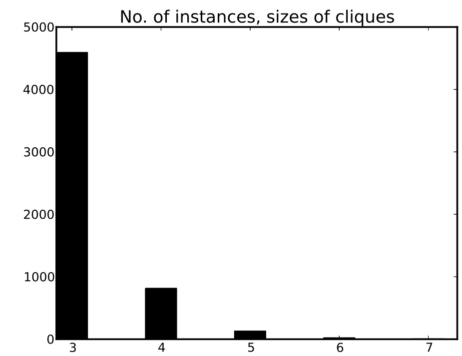


Figure 9: Number of instances and size of the cliques (only 23 cliques of size 6, and 6 cliques of size 7)

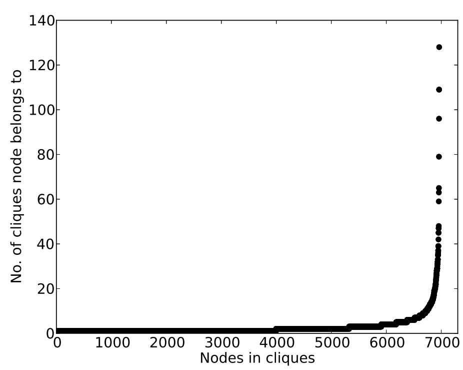


Figure 10: Number of cliques individual users were involved in.

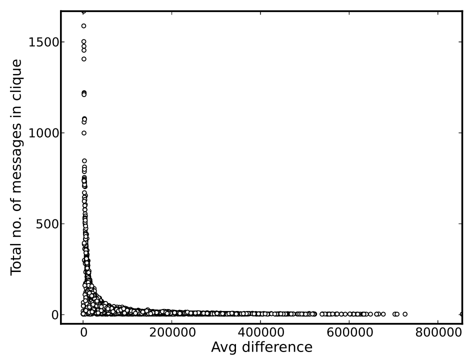


Figure 11: Average difference in seconds between two consecutive messages in clique versus total number of exchanges.

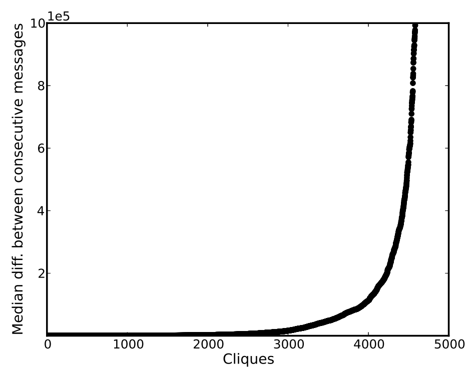


Figure 12: Sorted medians of differences in seconds between two consecutive messages for each clique.

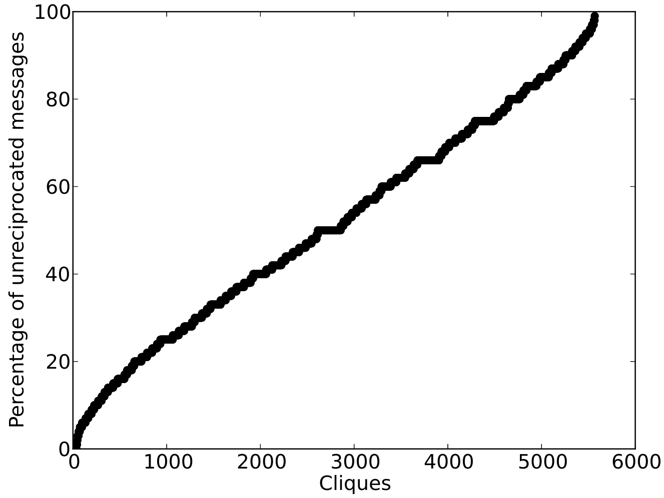


Figure 13: The percentage of 'unreciprocated' messages in cliques

ference of their in-degree and out-degree. We summed up the positive values[3] and to normalise, we divided by the total number of exchanged messages. In this way, we obtained a percentage of 'unreciprocated' messages, where reciprocity is not toward a sender but toward a whole group. We show the numbers over all cliques in Fig. 13. We see that all values are represented, i.e. that some cliques are almost completely dominated by certain members, but also that in some of the cliques every individual receives and sends a similar number of tweets.

Finally, we looked at so-called 'floor-gaining' [3], i.e. how much input each user had over the course of a group exchange[4]. We compared the out-degree of each user within a clique, (remember that each 'clique' is directed multigraph) with the mean number of edges $r = |n_\mathrm{E}|/|n_\mathrm{V}|$, where n_E is the total number of edges within the clique and n_V is the total number of vertices within the clique. In a 'round robin' group conversation, with balanced turn taking, each

[3]Clearly the number of sent and received messages within a group are equal, thus summing the differences between in- and out-degree over individual members in the group is by definition equal to zero.

[4]We argue that the action of tweeting in multiuser exchanges can be regarded as floor-gaining, since tweets with mentions can in principal be read by a wider audience than the group conversing.

Clique's size	No. of cliques	1 usr	2 usr	3 usr	4 usr	5 usr
3	4593	2045	2321	227	0	0
4	816	244	461	107	4	0
5	131	14	76	38	3	0
6	23	1	9	10	3	0
7	6	0	1	3	0	2

Table 1: A number of dominant users in clique

user would send out r messages, i.e. be responsible for an equal percentage $p = r * 100/e$ of the total number e of exchanged messages. For each clique size, we looked at how many users' representation were greater than or equal to p, i.e. those users who 'dominant' the conversation. In Table 1 below, we present the number of instances for each clique size and each number of dominant users. This shows that in most of the cliques, 2 users were responsible for the majority of communication, but a large proportion of exchanges were also dominated by a single user. However in about 6% of all cliques, 3, 4 or 5 users were dominating, confirming that Twitter is used for multi-user conversations and not just pairwise conversations.

5 Conclusions

We looked at conversations in Twitter, based on the underlying structure and timings in approximately 4 million UK tweets with mentions over a period of 28 days. To make use of graph algorithms, we structured the data as a multigraph. We proposed a simple method of identifying conversations between pairs of users, based on a time-threshold on the time-to-next tweet, and found evidence that a threshold of 9hrs gives a good indication of distinct conversations. We observed that the conversations detected using this method appeared to be balanced, meaning that each party involved contributed approximately equally to the conversation. This was not the case within more general interactions, in which one agent typically contributed a significant amount more than the other.

Although finding cliques in graphs is a computationally difficult problem, because of the sparsity of interactions patterns within the data-set, extracting multi-user exchanges (defined by the cliques) was feasible and relatively fast. We were able to find all cliques within the graph, up to a maximum of

7 users. While most of those exchanges were fast-paced i.e. had relatively small time gaps between messages, some were more relaxed. We also found that the number of messages in multi-user exchanges was inverse to the average time difference between them. When looking into balance of multi-user conversations, we found that most exchanges are dominated by just one or two users, with evidence of only a small percentage of well-balanced group exchanges. Regarding the number of received and sent messages by each individual in a group, we found that all kind of exchanges take place, i.e. groups with all different percentages of unreciprocated messages from 0 to 100% could be found.

Further work needs to be done using content information to explore how topics flow through multi-user exchange and if there is any relationship between time-differences between messages and topic. We hope that the insights gained from our analysis could help to develop an understanding of the mechanisms and dynamics of Twitter conversations, with potential scope for generating models of micro-blogging behaviour.

6 Acknowledgements

This work is partially funded by the RCUK Digital Economy programme via EPSRC grant EP/G065802/1 'The Horizon Hub' and EPSRC MOLTEN EP/I016031/1. We would like to thank Datasift for the provision of the data analysed and to Colin Singleton from Counting Lab Ltd. for pre-processing the data and providing useful feedback and discussion.

References

[1] J. Bollen, H. Mao, and Z. Xiao-Jun. Twitter mood predicts the stock market. *Journal of Computational Science*, 2:1–8, 2011.

[2] C. Bron and J. Kerbosch. Algorithm 457: finding all cliques of an undirected graph. *Commun. ACM*, 16(9):575–577, Sept. 1973.

[3] C. Edelsky. Who's got the floor? *Language in Society*, 10:383–421, 1981.

[4] S. Golder and M. W. Macy. Diurnal and seasonal mood tracks work sleep and daylength across diverse cultures. *Science*, 333:1878–1881, 2011.

[5] D. Graham. Twitter: value-added conversation is more important than broadcasts, rts. memeburn, August 2012.

[6] A. Hagberg, D. Schult, and P. Swart. Exploring network structure, dynamics, and function using networkx. In *Proceedings of the 7th Python in Science Conference (SciPy2008), Passadena, CA USA*, pages 11–15, August 2008.

[7] S. Herring. Computer-mediated discourse. In D. Tannen, D. Schiffrin, and H. Hamilton, editors, *Handbook of Discourse Analysis*. Oxford: Blackwell, 2001.

[8] C. Honeycutt and S. C. Herring. Beyond microblogging: Conversation and collaboration via twitter. In *Proceedings of the Forty-Second Hawai'i International Conference on System Sciences, Los Alamitos*, 2009.

[9] G. Leboff. Marketing is a conversation? oh yeah ...about what? The Marketing Donut, October 2011.

[10] E. Nuutila and E. Soisalon-Soinen. On finding the strongly connected components in a directed graph. *Information Processing Letters*, 49(1):9–14, 1994.

[11] D. Quercia, L. Capra, and J. Crowcroft. The social world of twitter: Topics, geography, and emotions. In *The 6th international AAAI Conference on weblogs and social media, Dublin*, 2012.

[12] A. Ritter, C. Cherry, and B. Dolan. Unsupervised modeling of twitter conversations. In *HLT-NAACL 2010*, 2010.

[13] J. Sidnell. *Conversation Analysis: An Introduction*. Blackwell, 2010.

[14] D. Starkey. Some signals and rules for taking speaking turns in conversations. *Journal of Personality and Social Psychology*, 23(2):283–292, 1972.

[15] R. E. Tarjan. Depth-first search and linear graph algorithms. *SIAM J. Comput.*, 1(2):146–160, 1972.

[16] K. Thomas, C. Grier, V. Paxson, and D. Song. Suspended accounts in retrospect: an analysis of twitter spam. In *Proceedings of the 2011*

ACM SIGCOMM conference on Internet measurement conference, pages 243–258. ACM, 2011.

[17] S. Wu, J. Hofman, W. Mason, and D. Watts. Who says what to whom on twitter. In *WWW, 2011*, 2011.

[18] S. Wu, C. Tan, J. Kleinberg, and M. Macy. Does bad news go away faster? In *5th International AAAI Conference on Weblogs and Social Media, 2011*, 2011.