A Characterisation of Mathematics Question-and-Answer Networks on Twitter

Juul Romkema
University of Twente
P.O. Box 217, 7500AE Enschede
The Netherlands
j.f.romkema@student.utwente.nl

ABSTRACT
Twitter is one of the most popular microblogging platforms these days. People use the platform to communicate, share information or to ask questions. Therefore, it will be no surprise that it is also used for educational purposes. This paper focuses on the use of Twitter for mathematics-related questions. It aims to understand the structure of this network as well as identify the extent to which questions are answered for new participants. Using this network, nodes are filtered based on whether they are friends who respond versus random users who respond. This will result in two network graphs. These graphs will be compared using topological characteristics. Subsequently, this paper will compare its outcomes with previous research on the Twitter community. The results suggest that the probability a mathematics-related question gets answered is small, however, it also suggest no difference in the probability for receiving an answer from a stranger or from a friend.

Keywords
Mathematics, Twitter, Topology, Hashtagged Q&A, Network Theory

1. INTRODUCTION
Social media platforms have changed the way we communicate. These platforms enable their users to share their opinion with people from all over the world and thereby create a wider audience for their views [5]. Several of these sites are microblogging sites. One of the most popular examples is Twitter. Twitter has 310 million monthly active users which send approximately 6000 tweets per second [17]. All these tweets have at least one similarity: they are limited to 140 characters. Photos and GIF images are not included in this limit although these attachments can contain additional information [19]. Social media platforms are especially useful for asking questions given the large potential reach of platform content. According to a survey amongst 624 Twitter users, 50.6% have used their status in the past to ask a question [11]. This makes social media platforms interesting for educational purposes. In 2012, 59% of the students declared to use their social media networks for educational matters [14]. The advantage of using Twitter would be the increase in engagement, the opportunity to network and collaborate, and the access to experts [10], compared to more traditional ways of asking questions. Another advantage is the ability to contact every Twitter user directly by using the @ sign. By using @username, the user receives a notification of an incoming message. He can then decide whether he answers or not. In this way, it is possible to ask questions directly to domain experts that are outside one’s network. An example of a conversation on Twitter is shown in Figure 1.

Another important feature of Twitter is the fact that the platform enables users to target a more specific audience by the use of hashtags. A hashtag is placed in front of a keyword to emphasise the purpose of a tweet. Clicking on a hashtag shows a list which contains other tweets tagged with the same keyword. In this way, a hashtagged Q&A can arise in which Twitter users can participate. An example of how hashtags are used is also shown in Figure 1.

Our dataset consists of mathematics-related questions on Twitter written in English. In this paper, we characterise the mathematics question-and-answer community on Twitter. We aim to identify the probability that one receives an answer to a question. Moreover, we also aim to determine the usefulness of the community for new members. To make sure the tweets are topic related, the dataset is composed of tweets which contain keywords that are often used in the mathematics community. A distribution will be made to show the chances that a question receives an answer. Next to this, we will also draw two network graphs: a static graph of who follows whom, which we call a mathematics friendship network, and a dynamic
graph of who responded to whom, which we call a mathematics activity network. Drawing these graphs provides a visual overview of the differences between the networks. To make these differences concrete we will look at the characteristics of both graphs. To be precise, the topological characteristics, which are the connections between users who do not change when the physical form of the network does change [3]. These characteristics will help us to characterise the friendship and activity network and will enable us to compare the graphs. From this comparison, an overview will be drawn.

1.1 Research Questions
These objectives result in the following research questions:

RQ1 How can tweets about mathematics be classified as questions and answers?
RQ2 What is the distribution of questions and answers tweets about mathematics?
RQ3 What are the topological characteristics of the mathematics activity network on Twitter?
RQ4 What are the topological characteristics of the mathematics friendship network on Twitter?
RQ5 What are the quantified differences between the topological characteristics of both graphs. To be precise, the characteristics will help us to characterise the friendship and activity network and will enable us to compare the graphs. From this comparison, an overview will be drawn.

2. BACKGROUND
In this paper, we are especially interested in the hashtagged Q&As about mathematics-related subjects. Therefore, first, a number of hashtags will be explained. Then, network theory for the identification of topological characteristics is discussed.

2.1 Mathematical hashtags
A hashtag can be obvious, but most of the time finding a proper hashtag can be hard. Different hashtags are used for different purposes. For example, #math is used by Twitter users to emphasise that their tweet is mathematics related. British English would rather use #maths, since they include a ‘s’ at the end of the word ‘math’ whereas the other forms of English do not. However, both of these hashtags are often used in a rather passive way and the user is only temporarily interested in mathematics.

Then there are also hashtags from people that are often used to refer to articles and other mathematics related material. Good examples are #mtbos, which stands for ‘math Twitter blog-o-sphere’, and #teammaths. In both cases the tweets include mathematics-related content and are aimed to start a discussion.

Furthermore, there are also broader hashtags, like #mathschat. This hashtag is used for conversations about mathematics and almost everyone interested in the topic, from parents to teachers, uses it. And then there is also a hashtag that is in line with #math in terms of audience, which is #mathematics. All the before mentioned hashtags can also be found in Table 1.

2.2 Network theory
The network behind Twitter is a directed graph, since user A can follow user B, but this does not mean that user B follows user A. In this paper two graphs will be formed: an activity graph and a friendship graph. Both have in common that their tweets are collected based on the hashtags mentioned in Table 1. From the collected tweets an selection of question and their answers is made. This selection is the basis for the activity graph, where a relationship is defined by the activity between two users. An arrow will be pointed from the respondent towards the person who asked the question. In case of the friendship network, besides the activity also the friendship relationship will play a role. In other words, in case of the friendship graph, an arrow will only be drawn when the respondent answered a question of the asker and the respondent follows the asker.

In order to gain more insight into both networks, a number of topological characteristics will be identified.

• Size of the network. The size of a network graph will be determined by means of the number of users N in each network. With the aid of this number it is possible to determine how many of the Twitter users are included in our mathematics question-and-answer network.

• Indegree and outdegree. The indegree and outdegree are determined by the degree of incoming edges k_i as well as the degree of outgoing edges k_o. These properties are also known as indegree and outdegree and from both it is possible to calculate the average degree of the network: \( \langle k \rangle \) and \( \langle k_o \rangle \). Based on both in- and outdegree it is possible to say something about the big influencers in both networks.

• Connectedness. The connectedness of a graph shows whether the community acts as one or whether there are multiple smaller components in the network connected by a few key users. Four categories of connectedness exist [4];
  1) Complete graph: all users are interconnected,
  2) Giant component: most users are interconnected,
  3) Weakly connected component: there is a path from every user to another ignoring the direction of the relationship,
  4) Strongly connected component: there is a path from every user to another taking into account the direction of the relationship.

• Diameter. The diameter of a network is determined by the longest path amongst all shortest paths. The diameter can be identified by looking whether the followers of user A are also interconnected. The diameter makes it possible to say something about the connectedness of the networks and its communities.

• Reciprocity. The reciprocity of a network tells something about the involvement of users in a network. It does so, by looking whether users feel obligated to repay a favour[2]. In this paper two kinds of reciprocity will be determined:

<table>
<thead>
<tr>
<th>Table 1. Mathematics-related hashtags on Twitter</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hashtag (#)</td>
</tr>
<tr>
<td>#mathstlp / #mathtlp</td>
</tr>
<tr>
<td>#mathscpdcchat / #mathcpcdcchat</td>
</tr>
<tr>
<td>#mtbos</td>
</tr>
<tr>
<td>#teammaths</td>
</tr>
<tr>
<td>#mathschat / #mathchat</td>
</tr>
<tr>
<td>#mathematics</td>
</tr>
<tr>
<td>#mats / #math</td>
</tr>
</tbody>
</table>

There are also hashtags that are used in a more regular and active way. An example would be #mathstlp, meaning ‘maths twitter lesson plan’, which mainly focuses on teachers sharing their knowledge. They do this during weekly organised sessions. Another hashtag used by teachers to prepare their lesson is #mathscpdcchat. This hashtag organises the lesson and is #mathstlp continuing professional development chat’. Teachers using this hashtag organise weekly sessions for each other, to improve their knowledge.
1) Friendship reciprocity. The friendship reciprocity is determined by the relationship between users in a graph G. For instance, we have responders and askers. First, the number of times the respondents, defined as A, and the askers, defined as B, follow each other is calculated as \( \#(A \leftrightarrow B) \). Then, the number of times only the respondents follow the askers and the askers do not return this favour is calculated, defined as \( \#(A \rightarrow B) \). The friendship reciprocity can subsequently be calculated using the following formula:

\[
FR(G) = \frac{\#(A \leftrightarrow B)}{\#(A \rightarrow B) + \#(A \leftrightarrow B)}
\]

2) Reply reciprocity. The reply reciprocity is determined by the activity relationship between users in a graph G. For this type of reciprocity only users that asked and responded to a question are relevant. If we take A and B from the previous example, then we want to know the amount of users that have asked questions and responded to questions, this is defined as \( \#(A \leftarrow * \text{ and } * \rightarrow B) \). Next to this, we also want to know how many users have answered each other’s questions, defined as \( \#(A \rightarrow B \text{ and } B \leftarrow A) \). This results in the following formula to calculate reply reciprocity:

\[
RR(G) = \frac{\#(A \rightarrow B \text{ and } B \leftarrow A)}{\#(A \leftarrow * \text{ and } * \rightarrow B)}
\]

All these types of topological characteristics help to characterise the mathematics question-and-answer activity and friendship network and are useful for a quantitative comparison. As discussed, two network graphs will be drawn in this paper. These will subsequently be compared using the characteristics as described above. Besides that, also our findings will also be compared with the findings of other research on topological characteristics of Twitter networks.

3. RELATED WORK

Prior research has given us some insight into online question-answering on Twitter and the topology of Twitter. A lot of people use social networks like Twitter to ask questions to other users. According to a study by Paul et al. who analysed 1152 questions on Twitter, only 18.7% received a response, with an average of 2.9 responses per question [15]. However, their research focused on the number of responses a user received based on their status, without emphasis on a certain area or the use of smart hashtags. In other words, chances are reduced that a tweet reaches the right audience and thereby the chances of getting a response are smaller. Rzeszotarski et al. argue that most replies are from within a user’s immediate follower network, contradicting the belief that everyone contributes to these hashtagged Q&As [16]. Another way of reaching users from outside one’s network was through direct contact which can be accomplished by using the @ symbol. A study by Nichols and Kang identified that 42% of the 1159 Twitter participants under study responded to direct questions from strangers [13]. With respect to the topology, we found papers that researched the topological characteristics of Twitter as a whole. A research, by Myers et al., into the structure of the Twitter follow graph collected a snapshot from the second half of 2012 which includes 175 million active users [12]. They found that the indegree was best fit with a power-law distribution, whereas the outdegree was best fit by a log-normal distribution. This same research also declares that 68.7% of the active Twitter users are included in the largest strongly connected component and that 92.9% of all active users are included in the largest weakly connected component. Furthermore, they found a diameter of 4.05.

We also found another paper on topology of networks on Twitter. This paper, a study by Kwak et al., used a snapshot of Twitter from 2009 which included 41.7 billion user profiles [8]. They found out that Twitter has a non-power-law follower distribution and a short effective diameter of 4.8. The word ‘effective’ points to the fact that this diameter is the 90th percentile distance. Furthermore, they found that Twitter has a low reciprocity of 22.1%.

The presented previous research on this topic all focused on Twitter in general and not on a specific part. Other related work is mainly focused on specific target groups, rather than a specific topic. For example, Macià et al. conclude that teachers act as bridges between several communities on Twitter [9]. Although the finding is relevant, it falls outside of our scope, since we focus on topics discussed on Twitter rather than on target groups. We will be the first contribution to focus on online questioning with regards to mathematics. We will compare the findings from this paper with related work to see if the mathematics network behaves the same as the other networks on Twitter.

4. RESEARCH METHOD

In this section, an overview of the research process will be presented. A schematic overview of this process is shown in Figure 2.

![Figure 2. Schematic Overview of Research Process](image)

4.1 Data collection

The first step of the research process was the collection of users and their tweets. In this section the data collection is described. To find out how the mathematics question-and-answer network on Twitter looks like, 97,830 tweets were gathered using Twitter’s Search Application Programming Interface (API) “GET search/tweets”. This API allows its users to crawl through tweets sent over the past seven days with a specific search query [20]. For this case, tweets were searched that contained at least one of the hashtags listed in Table 1.

For the first part of the dataset tweets were gathered from the 27th of April 2017 till the 31st of May 2017. This was done by executing the code twice a week. The code would run from the 27th of April 2017 till the 31st of May 2017. This was done by executing the code twice a week. The code would run from the 27th of April 2017 till the 31st of May 2017. This was done by executing the code twice a week. The code would run from the 27th of April 2017 till the 31st of May 2017. This was done by executing the code twice a week.
were merged and duplicates were removed. This resulted in an initial dataset of 41,367 tweets. An overview of all phases and the number of tweets is shown in Figure 3.

Table 2. Tweet properties

<table>
<thead>
<tr>
<th>Property</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>date</td>
<td>Date of the posted tweet</td>
</tr>
<tr>
<td>time</td>
<td>Time of the posted tweet</td>
</tr>
<tr>
<td>text</td>
<td>The actual content a user wrote</td>
</tr>
<tr>
<td>user_id</td>
<td>Unique number of the user</td>
</tr>
<tr>
<td>tweet_id</td>
<td>Unique number of the tweet</td>
</tr>
<tr>
<td>reply_to_user_id</td>
<td>Unique number of the user to whom is replied</td>
</tr>
<tr>
<td>reply_to_tweet_id</td>
<td>Unique number of the tweet that is replied to</td>
</tr>
<tr>
<td>hashtags</td>
<td>The hashtags which appear in the text</td>
</tr>
</tbody>
</table>

For the second part of the dataset, a list of users was created. This list of users contains users who have tweeted themselves, as well as users who are mentioned in the collected tweets. From this list, the 200 most recent tweets of a user’s timeline were requested. This was done using Twitter’s Rest API “GET statuses/user_timeline”. Next, all tweets that were retweets were disregarded as well as the ones that did not include one of the hashtags listed in Table 1. From the remaining tweets, all properties were requested again and these were written to a dataframe. In some cases, no tweets were found or the user had set his timeline settings to private. In this case, all property fields were set to ‘Not found’ or ‘Private’. In this way, the user was included in the list of users, which prevents that the code would identify the same user multiple times. Then, in the newly collected tweets, the algorithm would search again for users who were not in the list yet and would crawl through their timelines as well. This process is often called snowballing, since more and more users stick to the dataset. This snowballing was repeated until no new users were discovered in the dataset. This gave us a dataset of 66,316 tweets. From this dataset, we filtered out all the ‘Not found’ and ‘Private’ rows and merged the remaining rows with the first part of the dataset. After filtering out all duplicates, this resulted in a dataset of 96,616 tweets.

Up to now, only tweets with one of the specified hashtags were gathered. For the last step of data collection, again a list of users is created. This time the algorithm does not search for a tweet containing one of the hashtags, but rather it searches for tweets that reply to tweets that are already in the dataset. This last round of data gathering provided another 1,212 tweets. These tweets were added to the other dataset which resulted in our final dataset of 97,830 tweets. The data from the tweets in the final dataset ranges from March 2011 till June 2017.

4.2 Data filtering

The second step in the research process is to filter the data to create the mathematics question-and-answer network. This process starts by filtering out the tweets that contain at least one question. This is done by searching for tweets that contain a question mark (?) . Unfortunately, a lot of tweets contain a question mark, but do not contain a question. For example, this tweet:


To test the accuracy of this filtering method, a test set was composed of 500 tweets. The tweets were divided into two groups: questions and others. The first group consisted of 76 tweets, the second of 424 tweets. The filtering method predicted the classification of 476 tweets correctly. A complete overview of the outcome is shown in the confusion matrix in Table 3.

<table>
<thead>
<tr>
<th>Questions</th>
<th>True</th>
<th>False</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>True</td>
<td>64</td>
<td>12</td>
<td>76</td>
</tr>
<tr>
<td>False</td>
<td>12</td>
<td>412</td>
<td>424</td>
</tr>
<tr>
<td>Total</td>
<td>76</td>
<td>424</td>
<td>500</td>
</tr>
</tbody>
</table>

The filtering method misidentified 24 of the tweets, of which 12 are false positives and 12 are false negatives. Based on the confusion matrix in Table 3, the accuracy of the filtering method can be calculated. The accuracy will be determined based on the precision as well as the recall. These show that the accuracy is approximately 84.2%.

\[
\text{Precision} = \frac{TP}{TP + FP} = \frac{64}{12 + 64} = \frac{64}{76} \approx 0.842
\]

\[
\text{Recall} = \frac{TP}{FN + TP} = \frac{64}{12 + 64} = \frac{64}{76} \approx 0.842
\]

Then, when all questions were filtered out, a list of in_reply_to_tweet_id of every tweet in the large dataset is requested.
This property tells whether a tweet is in reply to another tweet. The property will be empty if it is not a reply; otherwise this property will contain a tweet_id of the tweet where is replied to. At the same time, a list of tweet_ids will be composed from the tweets that were identified as questions. Then the code will loop through both lists to find whether the list of in_reply_to_tweet_ids contains id's that are also mentioned in the tweet_id list. If rows are found which satisfy this requirement, the properties of these rows will be added to the filtered question since these are the answers to the selected questions. This resulted selection of 10,155 tweets in total. Subsequently, every conversation thus question and answers, gets a conversation number to make it easy to identify which tweets belong together. Next to this, also the friendship between the users involved in the tweet was identified. This was done with the use of Twitter’s REST API “GET friendships/show”. Most questions do not include a user mention, since the question is intended to be read by everyone. Because there is no user mentioned in these tweets, there will also be no relationship. However, answers always have at least one other user included, since this other user is the one who asked the question. For the answers and the question that do contain another user, the user_id of both users was entered into the API, after which the friendship properties between both users were replied. From this information, we extracted whether user A follows user B, defined as following, and whether user A is followed by B, defined as followed_by.

### 4.3 Data visualisation

The next step in the process is to visually represent the data. There will be two ways of data visualisation: the first one will show the distributions of question and answers in the dataset and the second will show the activity and friendship network graphs.

#### 4.3.1 Question and answer distribution

This subsection will describe how the question and answer distribution is made. To make this distribution only a subsection of the dataset is needed, namely the answers to the questions. This subset contains 671 tweets. From these tweets can be subtracted who asked the question, defined as in_reply_to_tweet_id, as well as who answered the question, defined as user_id. If an user answered its own tweet, the in_reply_to_tweet_id and user_id are identical. All tweets in which these properties were identical were removed from the selection, which resulted in a selection of 575 answers.

Tweets that have the same value for in_reply_to_tweet_id were requested again. These tweets all replied to the same tweet; the one that contains the question. In other words, this enabled us to identify how many answers a question received. Based on the information about the question and answer distribution, a graph was made that shows the number of answers on the x-axis and on the y-axis it shows the number of questions.

#### 4.3.2 Network graphs

This subsection will explain how the network visualisations are made. For the visualisations of the networks an open-source platform called Gephi is used [7]. This platform is often used by data analysts for the visualisation and analysis of network graphs. It allows its users to change structures, shapes and colours in order to discover patterns and to draw graphs especially manipulated for their purposes.

Before the network visualisations can be made data needs to suit Gephi’s format. The platform works with a distinction between nodes and edges. Therefore, our dataset needs to be rewritten into the format of Gephi. In other words, the activity graph will be rewritten into graph_a = (nodes_a, edges_a) and the friendship graph will be rewritten into graph_f = (nodes_f, edges_f).

For the nodes, Gephi demands at least a column with ‘id’s. This list of id’s can be seen as the name of the nodes. In our case, these are all users who wrote a question that is answered as well as the users that responded to these questions. We left all other nodes, i.e. the users with unanswered questions, out of considerations. This was done for the purpose of graph clarity. The nodes that are disregarded are 89% of the question and answer dataset. These nodes do not belong to any networks and therefore have no added value. The drafted node file will be the same for the activity graph as well as the friendship graph, since the same users are involved in both cases. In other words, node_a is equal to node_f. This will result in users without any edges in the second case, however, these users are important, since we want to compare both graphs and therefore the number of nodes should remain the same.

However, in case of the edges there is a difference between the activity and friendship graph. For the edges Gephi wants to know the Source, Target, and Type. In other words, the node who points towards the other is the source, the other node is the target and the type specifies whether this relationship is directed. To determine the edges for the activity graph, the dataset with only answers was taken again and the user_id was set as source node and the in_reply_to_user_id was set as target node.

In case of the friendship graph, an extra check was included: ‘there should only be a link between nodes from which the tweeter follows the receiver’. In other words, if the respondent follows the asker then an edge should be included, otherwise not. Next to the nodes, Gephi also wants to know the kind of relationship that exists between nodes and whether the relationship is Directed or Undirected. Since the graphs are based on the responses, the kind of this relationship is always Directed.

After the files have been imported into Gephi and the nodes and edges are added, it is time to visualise the data. The sizes of the nodes are determined by the amount of the outgoing edges and ranges between 10 and 100. The outgoing edges are chosen, since the size will then represent the activity of a responding user in this network. Next to this, the layout of the network graphs is also altered. First, all nodes are placed into a circle by the “Fruchterman Reingold” method. Subsequently, the “Yifan Hu” method was used to pull this circle somewhat apart, in order to enable us to see the separate components. Next, also the shape of the edges was changed to “Default Straight”, after which both graphs were saved to a .png file and a .pdf file.

### 4.4 Topological characteristics

The last step of the process is to determine the topological characteristics of both the activity and friendship graphs. For calculation of the topological characteristics all questions and answers will be taken into account, thus also the questions without any answers.

The first characteristic to determine is the size of both networks. Both graphs are based on the same dataset and therefore have the same size. This size is based on the number of users in the question-and-answer dataset of 10,155 selected tweets.

The next characteristics to determine will be the in- and outdegree. This is done with the aid of the NetworkX library in Python [1]. This library can determine the in-
well as the outdegree for every node. This will be done for every node. However, there are a lot of tweets without answers. If the users who wrote these tweets are not involved in any other form of activity, these nodes will be excluded from the in- and outdegree distribution graphs. This is for the same reasons as in the Question and answer distribution section: visual clarity. In the end, two of these graphs will be generated: one for the activity graph and one of the friendship graph. Further, the connectedness of both graphs are determined. This is done with the use of Gephi. This platform also has built-in algorithms to determine the connectedness of a graph. More specifically, the platform is able to determine the weakly connected components (WCC), as well as the strongly connected components (SCC) of the graph. Based on the WCCs, the giant component can also be determined since this giant component will be the biggest WCC. Next to this, there will also be determined whether there is a complete graph. An complete graph only exists when all nodes in a graph are interconnected, however the probability this will happen in a question-and-answer network is not likely. As for the diameter of the graphs, this is also something Gephi can determine. One should choose whether the diameter should be calculated with directed edges or undirected edges. Based on this, the built-in algorithm is able to determine the diameter of both graphs. Since our networks are directed, we choose to calculate the directed diameter. The last characteristic is the reciprocity of the network graphs. In this paper two types of reciprocity will be identified. The first one, friendship reciprocity, will say something about the kind of relationship between respondents and askers. First it is determined how many of the user pairs, thus respondent and asker, are both following or not following each other and in how many cases one follows the other, but not the other way around. Subsequently, we will divide these numbers according to the formula given in the Background section. This number will be equal for both graphs, since the relationship between users does not change based on their activity. Next, the second type of reciprocity, reply reciprocity, will be calculated. This second type will concern the incentive of users to return the favour of replying, which is to determine how eager users are to answer a question from someone who answered one of their questions in the past. This will be done based on the retrieved answers. First it is determined how many of the user pairs, thus respondent and asker, have both answered each others question. Second, we will determine the number of users that have asked a question as well as responded to the question of another user. Subsequently, we will divide these numbers according to the formula given in the Background section.

5. RESULTS
The final dataset consists of 97,830 tweets in total. After filtering for questions and answers 10.39% of the dataset was left, which equals 10,155 tweets. From these questions and their responses, the unanswered questions were disregarded, which resulted in a selection of 1,033 tweets. Based on the distinction between question and answers, a question and answer distribution was made. This distribution is shown in Figure 4.

What stands out is that the vast majority of the answered questions have received only one answer. The number of tweets with questions decreases as the number of answers increases, with one exception of a tweet which received 11 replies. However, the fact that the majority of the questions retrieved only one reply could be explained by the fact that questions are often information requests. Information requests ask for a specific answer, in other words, if the question is already answered it is useless for another user to send the same reply. A good example would be:

- How do you calculate a field of a pyramid? #help #math.

Although multiple answers are possible, there is only one answer that is correct. Due to this type of information requests, it is explainable why most tweets receive only one reply.

![Distribution of answered questions and their answers](image4.png)

Figure 4. Distribution of answered questions and their answers

This selection of tweets with answered questions and their answers is also loaded into Gephi for the visual representation of the network. This selection consists of 501 users, who represent the nodes in both network graphs. The number of edges was determined by the relationship between the respondent and the asker. In the activity graph, edges were added between all repliers and their askers. This resulted in 606 edges. The activity network graph is shown in Figure 5. For the friendship graph, on the other hand, an edge was only included in case the replier followed the asker. This resulted in a decrease in the number of edges, since not all repliers followed their asker. The number of edges was reduced to 355; the graph of this network is shown in Figure 6.

After both networks were formed, the topological characteristics were identified, based on all questions, both answered and unanswered, and the corresponding answers. The values of all topological characteristics can be found in Table 4.

<table>
<thead>
<tr>
<th>Number of answers</th>
<th>Activity Graph</th>
<th>Friendship Graph</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>299</td>
<td>50</td>
</tr>
<tr>
<td>2</td>
<td>22</td>
<td>4</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

The size of the nodes in both network graphs is based on the outdegree of the nodes. Therefore, the distribution of the outdegree is important for the visualisation of the network graph, therefore we decided to also include the degree distribution of the activity graph and of the friendship graph. These can be found in Figure 7 and Figure 8.

5.1 Comparison of the network graphs
What stands out in the visualisation is that both network graphs have one component that is clearly the biggest and this component is surrounded by smaller components. It seems that the number of components with only two nodes does not change in the friendship graph compared to the activity graph. This may be caused by two things: 1) the two-node components were friends, or 2) the probability that an asker receives an answer from a stranger is just as likely as the probability that an asker receives an answer from a friend. However, from these visualisations it cannot be derived which of the two explanations is true.
To verify which of the two statements it might true, all two-node components in the activity network are identified. The activity network consists of 168 two-node components. We once again attempted identification of these specific components in the friendship network. In this network, we found 94 two-node components, suggesting that in about half of the two-node components the respondent follows the asker. This suggests that statement 2) would be probable.

In Table 4 the topological characteristics of both graphs are mentioned. It is very striking that the values found for every topological characteristic are close together. The major differences lie in the diameter and the giant component of the network graphs. These two characteristics tell something about the connectedness of a graph, as well as the communities in a graph, however, the other characteristics which tell more about the activity in a graph are about the same. This would suggest that the probability that a mathematics related question gets answered by a friend is approximately equal to the one that describes getting an answer from a stranger.

However, this seems to be contradicted by the degree distribution. Figure 7 and Figure 8 show an overview of the degree distributions in the activity network as well as in the friendship network. Since the numbers remain relatively high in Figure 8 compared to Figure 7, it seems that a user is more likely to receive a response from a friend. However, to be able to say something useful we will need to look at the numbers instead of the overviews.

To say something more about the degree distribution we took a sample of nine users by the indegree and a sample of nine users by the outdegree. The samples consist of three users with the highest degree, three users around the median degree, and three users with the lowest degrees. First, the distribution of ingoing edges was determined. In the activity network there were 245 ingoing edges against 159 ingoing edges in the friendship network. This would suggest that more than half of the ingoing edges remains. The sample confirms this image.

Then, the distribution of outgoing edges was determined. The activity network consisted of 329 outgoing edges, while the friendship network consisted of 211 outgoing edges.

These numbers also suggest that at least half of the questions are answered by strangers and also in this case the sample seems to confirm it.

Based on these two samples it seems that tweets are more likely to be answered by friends than by strangers, but that their probabilities do not differ substantially. However, the findings may attain limited generalisability due to being obtained from a sample. As such, the conclusion remains a conjecture.

Table 4. Topological characteristics of the activity graph and the friendship graph

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Activity</th>
<th>Friendship</th>
</tr>
</thead>
<tbody>
<tr>
<td>Size of network graph (number of nodes)</td>
<td>3282</td>
<td>3282</td>
</tr>
<tr>
<td>Average indegree with zeros</td>
<td>0.15</td>
<td>0.09</td>
</tr>
<tr>
<td>Average indegree without zeros</td>
<td>2.03</td>
<td>1.82</td>
</tr>
<tr>
<td>Average outdegree with zeros</td>
<td>0.15</td>
<td>0.09</td>
</tr>
<tr>
<td>Average outdegree without zeros</td>
<td>1.51</td>
<td>1.47</td>
</tr>
<tr>
<td>Complete graph (boolean)</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Giant component (number of nodes)</td>
<td>305</td>
<td>141</td>
</tr>
<tr>
<td>Strong connected components (number of components)</td>
<td>3263</td>
<td>3275</td>
</tr>
<tr>
<td>Weakly connected components (number of components)</td>
<td>2860</td>
<td>3024</td>
</tr>
<tr>
<td>Network diameter</td>
<td>14</td>
<td>8</td>
</tr>
<tr>
<td>Friendship reciprocity</td>
<td>0.74</td>
<td>0.74</td>
</tr>
<tr>
<td>Reply reciprocity</td>
<td>0.08</td>
<td>0.11</td>
</tr>
</tbody>
</table>

5.2 Comparison with previous research

In this section the results of this paper will be compared with previous research to put results into perspective. This previous research is also discussed in the Related work section.
Another characteristic that has been calculated during this paper is the in- and outdegree of both networks. Myers et al. found a power-law distribution for the ingoing edges and a log-normal distribution for the outgoing edges. In Figure 7 and Figure 8 the in- and outdegree of both our network is shown along with a power-law trendline. The indegree, as well as the outdegree, seem to follow this trendline and are therefore probably power-law distributed.

Myers et al. also mention the number of users included in the biggest weakly connected component, as well as the number of users included in their biggest strongly connected component. In their case, the numbers were 68.7% for the biggest SCC and 92.2% for the biggest WCC. In this paper the numbers were much lower. The biggest SCC in the activity graph consisted of 16 users which is 0.14% of all users and the biggest WCC in this graph consisted of 305 users which is 2.72% of all users. The numbers for the friendship network are even lower with a biggest SCC of 4 users, 0.04%, and a biggest WCC of 141 users, 1.26%. These numbers clearly indicate a big difference between the connectedness of the networks of our research and those of other researchers.

The diameters found in previous studies were 4.05 [12] and 4.8 [8], respectively. These diameters are smaller than the ones we found. In the activity graph the diameter is 8 and in the friendship graph the diameter is 14.

Lastly, Kwak et al. found a reciprocity of 22.8% [8]. The reciprocity they discuss is comparable with our friendship reciprocity, which is equal to 74%. The fact that our friendship reciprocity is many times higher, is explainable by the fact that our network is unlikely to contain celebrities. Celebrities can be seen as outliers and are often the users who are followed by many people. In other words, celebrities can provide a distorted image. The lack of celebrities in our dataset can be one of the reasons for the big difference in reciprocity. We think the fact that most of our topological characteristics differ so much with previous research is related to the size of our dataset. Because our dataset is relatively small, probably a lot of users who belong to the mathematics question-and-answer network are not included in this dataset. Moreover, the previous research is from 2009 and 2012 which is relatively old considering the speed in which social networks change.

6. DISCUSSION & FUTURE WORK

A couple of factors have possibly influenced the results of this research. In this section these factors will be discussed as well as the reasoning behind it.

The first factor is regarding the size of the dataset. Due to time constraints and limits in Twitter’s API the dataset was relatively small. This ensures that nodes with unusual properties have more influence than they would have in an ideal situation. By unusual properties we mean, for example, that a node has an unusual high degree distribution or a question that has received an unusual high amount of replies. In case of a bigger dataset, these outliers are compensated more by the other nodes. The fact that our dataset is so small also makes it hard to compare it to previous research.

Another factor that might have influenced the results was the fact that no real natural language processing was used. The classifier built had trouble with recognising the difference between a reference to an article and a real question. Therefore, a filtering algorithm was made for this situation, however, this filtering method is most likely biased. The filtering method is based on sequences that often appeared in the tweets, however it is quite possible that the test set is not a good representation of reality. Its consequence would be that questions are wrongly identified as the ‘others’ category and that non-questions will be classified as questions.

Another factor that influences the results is the fact that images and links to websites are not considered in this research. Images cannot be included in a database and should therefore be processed by hand. Due to time constraints this did not happen. However, a lot of the images in mathematics tweets do contain a question to solve, for example, a puzzle. In other words, it could have led to more questions and answers. With regard to the links to websites, it was difficult to take them into consideration since URLs were also part of the filtering method to identify questions. Next to this, these URLs needed to be opened by hand, which again was not possible due to time constraints.

Yet another factor that might have influenced the results is the fact that we never looked into the results of every hashtag separately. Therefore, there is a realistic possibility that a couple of hashtags distorted the results. This is
especially realistic if one considers that the hashtags are relatively different. The last factor that might have influenced the results is the number of user mentions taken into account. The code was programmed to retrieve the user_id of a user who was mentioned in a tweet, however this method only retrieved the first user. In some cases, multiple users were mentioned in a tweet, however, except for the first one, the others were not taken into account. In case these users would also have been retrieved, the dataset could have been larger.

All above stated factors may have caused a smaller dataset and therefore another representation of the mathematics question-and-answer network on Twitter. Future research should aim to take these factors into account.

7. CONCLUSION
The dataset used in this paper indicates that around 10% of all mathematics tweets on Twitter are questions and related answers. However, the fact that only 4% of the questions is answered suggests that asking questions on Twitter is not very useful. The tweets that were answered were answered by friends as well as by strangers. Surprisingly, the probabilities of getting a response from a friend or from a stranger were not far apart. In other words, the probability a tweet gets answered does not seem to depend heavily on the network of the asker. This gives an outsider just as much opportunity to get its tweet answered as an insider in the mathematics community. Regarding the topological characteristics the numbers from both activity and friendship graph were relatively close to each other, with the exception of the diameter and the largest component. When comparing the largest components it stands out that the largest component is clearly much smaller in the friendship graph. Furthermore, the diameter also differs by six, which is a lot.

Finally, if we would compare these numbers with Twitter overall, it stands out that the values are much smaller in the activity graph. Furthermore, the diameter also differs by six, which is a lot.

Finally, if we would compare these numbers with Twitter overall, it stands out that the values are much smaller in the activity graph. Furthermore, the diameter also differs by six, which is a lot.

8. ACKNOWLEDGEMENTS
I would like to thank my supervisor D. Bucur for her guidance and extensive feedback. Furthermore, I would like to thank M. Habib for his guiding through the data filtering process. Lastly, I would also like to thank Guido Teunissen, Jan Thiemen Postema and Tycho Teesselink for peer reviewing my writings.

9. REFERENCES