2016 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)

Toward Understanding How Users Respond to Rumours in Social Media

Anh Dang[1], Michael Smit[1], Abidalrahman Moh'd[1], Rosane Minghim[2], and Evangelos Milios[1]

[1]{anh,amohd,eem}@cs.dal.ca, {mike.smit}@dal.ca, Dalhousie University, Halifax, NS, B3H 4R2, Canada
[2]rminghim@icmc.usp.br, University of São Paulo-USP, ICMC, São Carlos, Brazil

***Abstract*—As the spread of rumours has been increasing every day in online social networks (OSNs), it is important to analyze and understand this phenomenon. Damage caused by the spread of rumours is difficult to handle without a full understanding of the dynamics behind it. One of the central steps of understanding rumour spread is to analyze who spread rumours online, why, and how. In this research, we focus on the steps *who* and *why* by describing, implementing, and evaluating an approach that studies whether or not a group of users is actively involved in rumour discussions, and assesses rumour-spreading personality types in OSNs. We implement this general approach using Reddit data, and demonstrate its use by determining which users engage with a recurring rumour, and analyzing their comments using qualitative methods. We find that we can reliably classify users into one of three categories: (1) "Generally support a false rumour", (2) "Generally refute a false rumour", or (3) "Generally joke about a false rumour". Combining text mining techniques, such as text classification, sentiment analysis, and social network analysis, we aim to identify and classify those rumour-spreading user categories automatically and provide a more holistic view of rumour spread in OSNs.**

I. Introduction

Online rumours, which are truth-unverifiable statements in online social networks (OSNs), are popularly spread in uncertain situations [1]. As billions of people and organizations are connecting with each other through social interactions, breaking news, and sports events, OSNs has become a popular source to share credible information [2]. As well as spreading credible information, OSNs can spread rumours [3], [4]. Problems like rumours going viral are not isolated and prompt the question of how to identify and limit the spread of rumours in OSNs. In an effort to constrain the spread of rumours, many researchers are trying to detect rumours [5], [6] and the original sources of rumours [7], [8]. However, little work has been done to understand who spreads rumours online and why.

People spread rumours for a variety of reasons. Bordia and DiFonzo [9] studied, from a psychological viewpoint, what motivates people to spread rumours in a social network. The authors identified three motivations for people to spread rumours: (1) fact-finding, (2) relationship-building, and (3) self-enhancement. People who are motivated by fact-finding are aiming to arrive at a valid and accurate understanding of rumours through a problem-solving process. In contrast, those motivated by relationship-building are simply interested in interacting with other people by sharing information about particular rumours. The same study pointed out that those with self-enhancement as motivation are either consciously or unconsciously simply spreading rumours. Researchers have tried to group users into the same group based on link predictions [10] and content characteristics [11]. In this work, we aim to automatically detect users based on how they interact with a rumour with regards to Bordia and DiFonzo's rumour-spreading theory.

As people have a tendency to believe misinformation, which are false rumours, and misinformation is more likely to be spread [12], we focus on studying whether users spreading misinformation in Reddit can be divided into one of the three categories derived from Bordia and DiFonzo's rumour-spread motivation theories: (1) "Generally support a false rumour" (self-enhancement), (2) "Generally refute a false rumour" (fact-finding), (3) or "Generally joke about a false rumour" (relationship-building). To achieve this goal, the proposed research collects rumour-related data from Reddit and applies text mining techniques and social network analysis to analyze and visualize users in those three categories.

To date, most of the work in this emerging area has been conducted to: detect rumours, limit the spread of rumours, and identify the source of rumours. However, in order to develop effective methods for rumour detection and prevention in OSNs, we first need to understand who spreads rumours online and why. This motivates us to propose the following research statements:

1) Based on user activities in Reddit, could we determine if there is a specific group of users that is greatly interested in discussing and spreading rumours?
2) Based on user activities in Reddit, could we determine if there is a rumour-spreading personality type in Reddit who, for example, "Generally supports a false rumour" (SUPPORT), "Generally refutes a false rumour" (REFUTE), or "Generally jokes about a false rumour" (JOKE)?
3) Will visualizing rumour spread in Reddit provide better insight into how users interact with rumours?

This paper makes the following contributions:

- Collecting and analysing user posting behaviours in Reddit about a specific rumour. Based on users' interaction, determine if there is a group of users that is actively spreading rumours.
- Using social network analysis, visual analytics, content analysis, and text mining techniques, the system classifies the active rumour-spreading users Bordia and DiFonzo rumour-spreading theory [9] into one of

IEEE/ACM ASONAM 2016, August 18-21 2016, San Francisco, CA, USA
978-1-5090-2846-7/16/$31.00 ©2016 IEEE

the three categories: (1) SUPPORT, (2) REFUTE, and (3) JOKE.

- The experimental results using text mining techniques confirm and support our approach.

This paper has the following structure: Section II reviews related work, Section III describes how we collect the data, Section IV describes the methodology, and Section V analyzes and discusses the results.

II. Related Work

Previous work in this area is concentrated in three main areas: mining online social networks, rumour analysis, and visualizing rumour spread in OSNs. It is important to highlight that the research focuses on rumour spreading in social media. It is implemented for Reddit data, and illustrated with the "Obama is a Muslim" rumour.

A. *Mining Online Social Networks*

Modern OSNs produce vast amounts of user-generated content. Analyzing content at this scale requires algorithmic support, typically in the form of data mining. Falkowski et al. [13] used statistical analysis and visualizing OSNs to study the dynamics and evolution of subgroups in communities. The authors proposed different community similarity measures and grouped similar communities into the same cluster, and later visualized community-clustering results to analyze their dynamics and evolution. The experiments showed that this method could detect the fluctuating nature of an online community. Liben-Nowell and Kleinberg [14] adopted knowledge from social network analysis (e.g., centrality features), graph theory (e.g., graph distance), and social sciences to gauge the effectiveness of network-proximity measures. Based on these measures, the authors tried to predict new interactions that would have a high probability of occurrence in the near future. Golbeck et al. [16] predicted the personality type of a user based on the user's Facebook profile. Dang et al. [17] uses text syntactic and semantic similarity to map related Tweets to users' profiles.

None of the studies examine Reddit data; Reddit is understudied in the social media research literature, despite being one of the most-visited social websites in the US.

B. *Rumour Analysis in Online Social Networks*

Within the field of social media research, there has been previous work focused on rumour analysis, using a variety of approaches (including data mining). In this research, we are only interested in using rumour-related memes to pinpoint which users are spreading or refuting rumours in OSNs.

The modern study of rumours dates back to 1944, in the work of Festinger et al. [18]. The authors studied the origin and spread of rumour in a specific neighbourhood community by intentionally starting rumours. After six months, intensive open-ended interviews with the residents in this neighbourhood about the rumours were recorded. The experiments found that not everyone who heard the rumour spread it further, and existing friendship connections between people increased the probability of the rumours being spread. Due to the intrinsic long-lasting nature of rumours and the difficulty in collecting rumour data, rumour analysis research experienced a lengthy hiatus until the popularity of OSNs in the 2000s.

In most OSNs, information is disseminated and stored permanently, so researchers are able to use the data to study rumours and their analysis more effectively. Marett and Joshi [19] investigated underlying motivations for posters and lurkers spreading information and rumours in a local online community. Posters are users that regularly post their experiences and stories in OSNs, while lurkers are users who only read the posts from other posters. The authors gathered posting data from a local university forum and conducted an online survey for both posters and lurkers in that community to understand why they spread rumours. The results showed that the intrinsic motivation, i.e., "the doing of an activity for its inherent satisfaction rather than for some separable consequence" [20], played a critical role in motivating posters to share information and rumours in this online community. One limitation of this approach is that it relied on self-reported responses from users to hypothesize why users spread rumours.

Recently, researchers have focused on using machine learning and the availability of big data in OSNs to study the spread of rumours and detect them automatically. Shah and Zaman [8] built a probabilistic model graph based on network structure and rumour-infected users. This model provides a rumour centrality score for each node in the graph, and the node with the highest rumour centrality score is the source of rumours. Qazvinian et al. [21] proposed a general supervised-learning framework to identify rumours in Twitter. Retrieved tweets were manually labeled as either being related to rumours or not. Based on this training set, the machine-learning framework classifies whether or not incoming tweets are about the rumours.

Although researchers have achieved some degree of success in detecting rumours and understanding their pattern, little work has been done to investigate who spreads rumours in OSNs and why. The closest work to our research is that of Buntain and Golbeck [16], who presented an automated method for identifying the "answer-person" role in Reddit based on user interactions. Users filling this role only respond to questions by other users and do not get involved (or have only limited involvement) in other discussions. They first manually analyzed data collected from Reddit to determine if this role exists. Next, they designed a feature set that characterizes this role and uses this feature set to classify more answer-person roles in the network. Our goal parallels the work of Buntain and Golbeck; our objective is instead to determine if rumour-spreading users exist in OSNs. To the best of our knowledge, no similar work has been done on studying rumour-spreading users in Reddit.

C. *Visualizing Rumours in Online Social Networks*

One of the most effective ways to study rumours in OSNs is to visualize the paths and patterns of the spread. Some recent scholarly and industry-led projects relied on visualizations to show how online rumours are spread. Ratkiewics et al. [22] developed *Truthy*, a supervised-learning visualization framework, to identify misleading political campaigns by collecting, analyzing, and visualizing messages through the Twitter network.

First, this framework detects any emerging memes which are a unit of information that can be spread from users to users in Twitter. Next, content, network and sentiment analysis are used to classify whether a meme is rumour-related. Finally, the path and pattern of rumours are visualized for further research. Similarly, The Guardian [23] visualized how rumours identified by reporters covering the story about the 2011 UK riots spread on Twitter by grouping related Twitter messages into the same cluster. Dang et al. [24] proposed RumourFlow, a visual analytics framework, which allows analysts to collect, analyze, and visualize rumour spread in Reddit by exploiting the use of social science theories, text mining techniques, information diffusion models, and sentiment analysis. In this work, we use visualizations, text mining techniques, and social network analysis to analyze and understand how rumour-spreading users interact with rumours and with other users.

III. Reddit Social Network

Reddit, which claims to be the front page of the Internet, is a social news website where users can actively participate in content creation. Registered users discuss a wide range of topics such as politics and world news every day. User-submitted content, called *submissions*, can be text content and direct links to other online content. Redditors can comment or vote (up-vote or down-vote) on each submission; these interactions determine the rank of the submission on the site. Redditors organize content into subcategories called subreddits. Every Reddit submission has the following elements:

- **Title:** The title summarizes the topic of the submission and is usually very short and concise.
- **Comments:** for each submission, users can post a comment that expresses their opinion about the submission; comments are organized hierarchically, so users can post comments on other comments. Users can also vote the comment up or down.
- **URL:** each submission may contain a link to an external source of information that is relevant to the submission.
- **Image:** each submission may also contain a link to an image to illustrate what the submission is about; a thumbnail is displayed on Reddit.

Reddit.com is ranked as one of the most visited sites globally. The massive amount of data disseminated through Reddit every day makes it an excellent tool for analyzing and detecting rumour-spreading users in social media. Although Twitter has been the most popular source for studying rumour spread in OSNs [21], [22], Reddit has made a few inroads into the world of analyzing rumour-related memes [25], [16]. While Twitter and Reddit do share some commonalities, they are different in important ways. Twitter primarily circulates news through known cycles (e.g., “follow” connections), whereas Reddit promotes a constant stream of new links to all users through a simple bookmarking interface. This makes Reddit an effective source for studying the spread of rumour-related memes in OSNs.

IV. Methodology

We describe our general approach to collecting, visualizing, and analyzing rumour-related data using Reddit as a specific implementation example.

A. Data Collection

To study the spread of rumours in Reddit, we need the following elements:

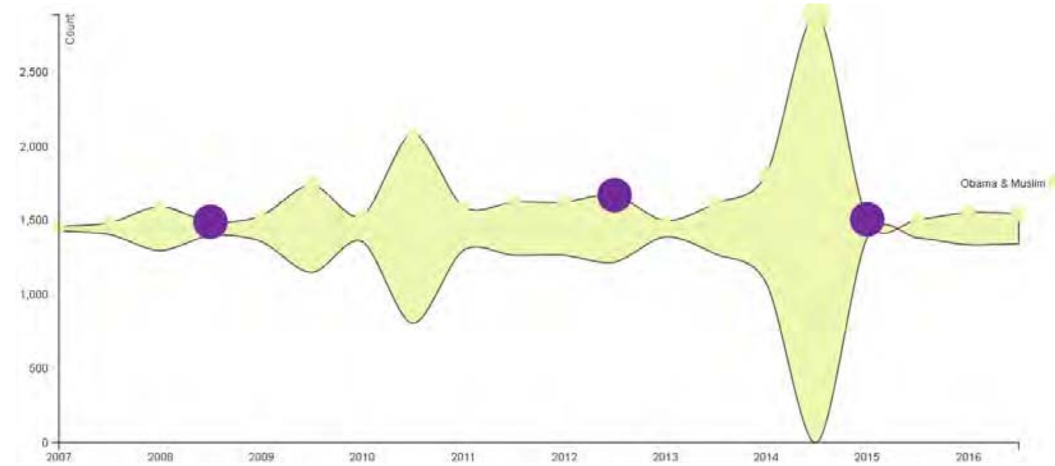
1) A rumour.
2) The truth about this rumour.
3) Posts about this rumour.

We adopt Snopes.com as a reliable source for collecting, confirming or disapproving rumours. Snopes is a website that collects memes, urban legends, and stories with unknown or uncertain origins. It provides a wide range of rumours, from politics, altered images, to real photos with fake stories, and even hoaxes. Each rumour is categorized as true, false, partly true, multiple truth-values, unverifiable, or legend. The editors of this website verify and provide evidence that could be used for debunking or confirming rumours. We also collect submissions and comments that are related to a specific rumour being discussed in Reddit. There may be one or more submissions for each specific rumour, so we have to create a generic query to capture all of the rumour-related submissions. Since no repository for a Reddit rumour dataset exists, we use the Reddit API and jReddit (a JAVA open-source project) to extract submissions, comments, and other data views, such as image or URL content, about a specific rumour using predefined regular expressions.

We used the rumour “Obama is a Muslim” in Reddit from 2007 to 2015 as our case study due to its persistence, popularity, and controversy. We automatically searched the keywords “Obama & Muslim” from Reddit and collected 195 submissions, 26,421 comments from 11,125 users, 85 submissions containing a URL, and 29 submissions containing an image. As our primary interest was in users that are actively involved in rumour spread, we removed users that engaged

TABLE I: Examples of submissions about the rumour “Obama is a Muslim”.

No.	Title	Date	No. Comments
1	People in Middle America believe that Obama is a Muslim	2007	234
2	Is Obama a Muslim? About.com poll: 57% Yes, 37% No, 10% Undecided. Let’s correct this.	2008	299
3	Do you think Mr. Obama is a Muslim or a Christian? ....I know, I know...	2009	28
4	Scientist asks why Americans believe Obama is a Muslim	2010	279
5	Iowa GOP Focus Group: Obama Is A Muslim	2011	169
6	Do you think Barack Obama is a muslim? Alabama Republicans: 45% say yes. Mississippi: 52%	2012	169
7	My Orthodox rabbi says President Obama is halachicly a Muslim...	2013	41
8	Proof that Obama is a Muslim!!!	2014	30
9	Poll: 54% of Republicans say that, "deep down," Obama is a Muslim	2015	2923



(1)

Fig. 1: Submissions regarding "Obama is a Muslim" over time. Each node represents a submission, and the nodes are visualized in ascending order of posted time while the y-axis represents the number of comments of each submission. The purple nodes represent the submissions that the user "kickstand" has commented on.

in fewer than 10 comments in these 195 submissions. This reduced the number of users to 163, and is the dataset used most frequently in this paper (note that we choose 10 as our threshold value to achieve a good representative sample of active users and to ensure statistical significance). Given our interest in not just assessing the existence of this group, but also in assessing if long-term participants in the conversation can be categorized into categories based on Bordia and DiFonzo's described motivations. Two judges review comments of each user based on their repeated comment patterns and categorize them into one of the three categories SUPPORT, REFUTE, and JOKE (with Kappa agreement score = 0.85).

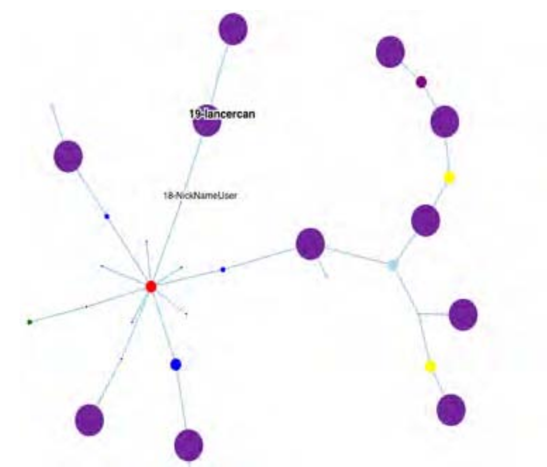
B. Data Visualization

For data visualization, this paper uses RumourFlow [17], a service-oriented visualization framework to collect and visualize rumour spread. All collected rumour data are provided to the visualization system through a JSON restful web service from a JAVA backend. For visualization, we adopt D3 and jQUERY to display rumour spread through a web-based application. The goals of this visualization framework are to provide a visual analytics tool for researchers and end users to explore different aspects of rumour spread in OSNs. It has two main views. The first view presents an overview of how rumours evolved over time as shown in Figure 1, and the secondary view describes how users interact with each submission about rumours by an egocentric network as shown in Figure 2. This framework also offers users easy access to search for a specific rumour in Reddit with their own keywords or for a specific user that comments about a rumour.

C. Approach to Analysis

After collecting the data, we adopted social network analysis, content analysis and text mining techniques, to analyze and visualize these contents.

Social network analysis (SNA) refers to the use of network theory for understanding social network data. Social networks



(2)

Fig. 2: User "lancercan" interaction graph about the rumour. A node represents a comment for a specific submission, and a node size displays how many times a user comments on a specific submission. The red node represents the original submission, and comments in the same level are nodes of the same color (e.g, the blue nodes represent comments to the red nodes). The purple nodes represent the comments of user "lancercan" for this submission, and suggest "lancercan" is a frequent participant in this discussion.

have been widely used since the early twentieth century to depict a certain community and how people in this community interact [26]. Because of the analogy between online social networks and the structure of social hierarchy and stratification, the study and analysis of social networks has played an important role in understanding how OSNs work. We focus on how users interact with other users about this rumour using our visualization tool to explore the data, particularly connections between users and a longitudinal assessment of the prevalence and re-appearance of the rumour over the 9 years in our dataset.

Content analysis is a qualitative method that examines the meaning of textual data manually to identify and assess themes and patterns. We focus on the characteristics and content quality of user posts in each category of user (SUPPORT, REFUTE, JOKE) and manually review each comment in all three categories to identify typical patterns and themes.

Finally, text mining techniques, such as data classification, visualization, and sentiment analysis are used to validate if the characteristics of each user group found from social network analysis and content analysis could be classified automatically.

All rumours have a beginning and an end. A rumour may be considered true at one point but is debunked as false at a later point. As a result, we try to capture all submissions about a rumour and visualize its evolution from its start to its end so that end users can discover all facets of a rumour life cycle. An example of each submission about the rumour "Obama is a Muslim" in each year from 2007 to 2015 is shown in Table I. An interesting observation is that the submission in 2015 still receives numerous comments from users. This suggests that the "Obama is a Muslim" rumour is still popular, even though it was first started in 2007.

V. Results

This section describes the analysis of the data, focusing on the highlights of the examination of the data using the visualization tool towards answering the research questions.

A. Rumour-discussing Users

TABLE II: Rumour-spreading users about the "Obama is a Muslim" rumour.

Rumour-spreading Users	User Count	Percentage
SUPPORT	8	4.9%
REFUTE	41	25.2%
JOKE	85	52.1%
OTHERS	33	20.2%

TABLE III: Examples of user comments in each category.

Category	Comments
SUPPORT	"He is a Muslim clearly."
REFUTE	"This is actually a good point. The radical conservative movement doesn't use language like the rest of the people. They don't say what they mean, or what they think is true. They say things to achieve the desired result. So, if they think saying Obama is a Muslim will damage him, by all means they will say that. They use "words that work"?"
JOKE	"?Eh you should come to the south and meet the people I have. Many people seriously believe he's Muslim. Many people also think men have less ribs than women despite that we know 100% it's not true. People are stupid."

First of all, we aim to determine if there is a group of users actively involved in rumour discussion and spread. Of the 11,000 users that have comments in the 195 submissions, 163 users have repeatedly interacted with one or more submissions by having 10 comments or more in those submissions. For example, how the user "lancercan" actively interacts with a submission about the topic "Obama is a Muslim" is shown in Figure 2. Another example shows how the user "kickstand" interacts with the submissions in the collected dataset in Figure 1. The use of stream and circles for visualizing time series graph has been used widely in the literature [27], [28]. This visualization helps to discover how a user is actively involved in discussing and spreading rumours. It shows that this user has repeatedly commented on this rumour since 2007 until 2015 in various submissions and years. These examples illustrate the larger group of users on Reddit that is very interested in discussing this topic for an extended period; the existence of this group is clear from the data.

A breakdown of the user-category dataset statistics is shown in Table II. The data demonstrates that most users either joked about this rumour or refuted it with a detailed explanation. Only a small portion of users supported this rumour. An example of user comments in each category is shown in Table III. Users in "OTHERS" categories seems to discuss related points with the rumour. Some of them discuss religion related topics.

B. Cross-Category User Interactions

We also counted the possible connections among users in the three categories (i.e., who replies to whom) to explore how users in different categories interact with each other. An example of each interaction between user groups is shown in Table IV and detailed statistics on how users in one group interact with users in another group are shown in Table V. In these two tables, each row represents how users in that category reply to users in the other three categories. There is not enough data about how users in the category SUPPORT interact with each other or with users in the other categories. It is clear that users in JOKE category tended to receive more responses from other users. Users in REFUTE and JOKE categories share many interactions between them. One possible reason for this finding is that Snopes debunked this false rumour in 2009, so people are inclined to refute or joke about it; there were very few supporters in our dataset. Furthermore, users in JOKE are more likely to respond to a comment of a user in SUPPORT or REFUTE and are also more likely to have connections to users in the same category. This may suggest that humour may play a significant role in why this rumour is so popular.

We also investigated how users in all three categories interacted with a specific submission about this rumour. We found that submissions that posted a link or an image to an external source perceived as reliable received much more attention and many more comments from users than a submission without a link or an image. This suggests that rumours with an image or a link from external sources perceived as reliable are more likely to be spread further [29].

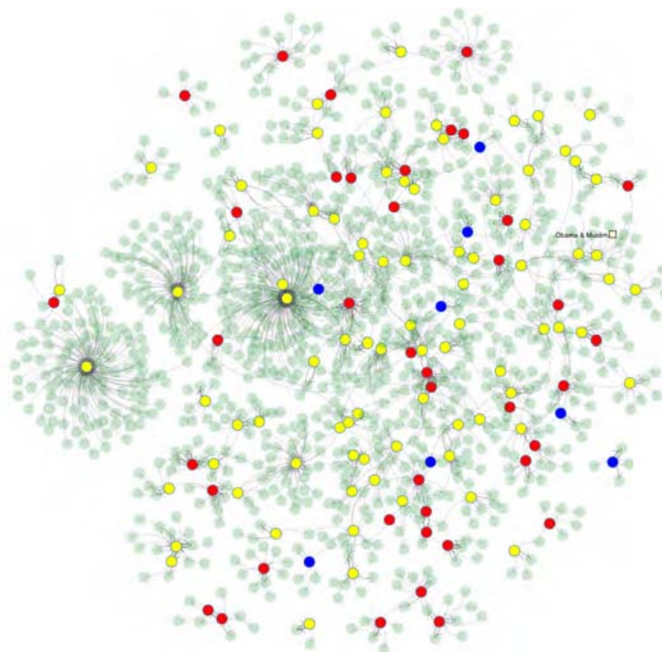
C. Content Analysis

Beyond understanding the users and their interactions, we also sought to analyze the textual content of submissions in each category. As most users that are actively engaged in this rumour do not believe it is true, we revisited the original dataset, which includes users who commented fewer than 10 times. We found that users in SUPPORT usually posted only one or two comments about this rumour. All of these comments were usually very short and had no back-up evidence or explanation. Here are a few examples:

"He's the kind of Muslim who?"
"I think he's a Muslim too"

TABLE V: Interactions between user groups; each row represents how frequently users in that category reply to comments or submissions in each of the three categories.

Number of connections	SUPPORT	REFUTE	JOKE
SUPPORT	N/A	N/A	N/A
REFUTE	25	5	4
JOKE	63	49	34



(3)

Fig. 3: An example of "who replies to whom" ground-truth user graph. Yellow nodes: users in JOKE category. Red nodes: users in REFUTE category. Blue nodes: users in SUPPORT category. Green nodes: users that have no more than 10 comments but has a connection to the users in one of the three categories.

"He's clearly trying too hard to not look like a Muslim. That makes it obvious that he is actually a Muslim."

For users in the REFUTE category, many comments were very thoughtful and provided in-depth explanations. Here are a few examples:

"Right, a politician would never lie or dissemble. If Obama says it, it must be true. I don't think Obama is a covert Muslim, but I wouldn't be surprised to learn that at some point in his life he was saying the [Shahadah.](http://en.wikipedia.org/wiki/Shahadah) His father was a Muslim before being an atheist. His mother ran off with a man to Indonesia and brought Barack Hussein with to the most populous Muslim nation. Barack doesn't strike me as a Muslim, but he may have real Muslim sympathies and may well have been exposed to Muslim indoctrination in his youth. Rejecting this possibility on the word of a lawyer and politician is your prerogative, but I prefer rational skepticism when it comes to politics."

"I don't care about any candidates religion unless they point it out as one of there qualifications for being elected to office. I can't remember Obama doing that except to deflect comments that he is a Muslim. Many republicans point out there adamant belief in Christianity and the belief that man was created by God, as stated in the First Book of The Bible: Genesis as a scientific fundamental. I cannot bring myself to vote for that type on nonsense. So I usually just waste my vote on a third party candidate."

Users in the category JOKE usually made a sarcastic comment or joke to refute this rumour. Here are a few examples:

"Mitt Romney's Birth Certificate. His Father was born in Mexico. Romney is just as 'foreign' as Obama is Kenyan or Muslim."

"Instead of convincing all those people they are wrong, we should just get Obama to convert to Islam."

In this instance, people are more prone to make a joke about it or refute it with a detailed explanation. Only a few people believe or support it.

D. Sentiment Analysis

In an online conversation, users' sentiment analysis has played a major role how this conversation becomes popular

TABLE IV: Examples of comment interactions between user groups.

Category	SUPPORT	REFUTE	JOKE
SUPPORT	N/A	N/A	N/A
REFUTE	**Rumour:** "Poll: 54% of Republicans say that, 'deep down', Obama is a Muslim". **Comment:** "Dam it to hell, I knew he was a Muslim!". **Response:** "I knew you are wrong."	**Rumour:** "Poll: 54% of Republicans say that, 'deep down', Obama is a Muslim" **Comment:** "Funny, because I suspect if he were a closeted anything, it'd be a closeted atheist." **Response**: "He's an atheist because he don't believe in god?"	**Rumour:** "Poll: 54% of Republicans say that, 'deep down', Obama is a Muslim" **Comment:** "Deep down, 54% of Republicans are idiots" **Response:** "I agree"
JOKE	**Rumour:** "Poll: 54% of Republicans say that, 'deep down', Obama is a Muslim". **Comment:** "He's clearly trying too hard to not look like a Muslim. That makes it obvious that he is actually a Muslim". **Response:** "Except, he will always look like a Muslim"	**Rumour:** "Poll: 54% of Republicans say that, 'deep down', Obama is a Muslim" **Comment:** "Deep down, 54% of Republicans are idiots." **Response**: "More proof that American voters have little or no memory."	**Rumour:** "Poll finds 23% of Texans think Obama is Muslim" **Comment:** "Poll finds 23% of Texans are idiots" **Response:** "I like to look at the positive side: 77% are not stupid"

and its topic evolution [30], [31]. Each comment was parsed into sentences and each sentence is assigned a sentiment score: "positive","negative", and "neutral" using OpenNLP. We apply the concepts of sentiment polarity and subjectivity of Zhang and Skiena [31] for each user category in our ground-truth dataset as follows:

$$polarity\_score = (p-n)/(p+n)$$

$$subjectivity\_score = (n+p)/N$$

where p is the number of positive statements, n is the number of negative statements, and N is the total number of statements (including neutral statements). Polarity score represents if a user category is associated with the entity positive or negative, while subjectivity score depicts how much sentiment a user category garner. The polarity and subjectivity scores for each user category are shown in Table VI. REFUTE users have the highest polarity and subjectivity scores, while SUPPORT users have the lowest polarity and subjectivity scores. This can be explained as this rumour was debunked by Snopes in 2009 as a false rumour.

TABLE VI: Polarity and Subjectivity Score of Each User Category.

Category	Polarity	Subjectivity
SUPPORT	0.484	0.680
REFUTE	0.747	0.753
JOKE	0.638	0.705

E. Classifying Rumour-spreading Users

Using the content analysis, we observe that content characteristics in each rumour-spreading user group has its own characteristics. As a result, in this section, we explore if we could determine the user rumour-posting behavior automatically based on its content. For each user that has more than 10 comments, we transform them using the TF-IDF vectors, which reflect how important a word is in a document or a corpus (stop words are removed). Each user is represented by a vector:

User = $\{T_1, T_2, ..., T_n\}$ where T_i is TF-IDF score of term i by the formula $tf-idf_{t,d}$:

$$tf-idf_{t,d} = tf_{t,d} \times idf_t = tf_{t,d} \times log(\frac{N}{df_t})$$

where t is the term, d is the comment that has term i, and N is the number of users (documents).

After transforming each user comment data into a TF-IDF vector, we apply NaivesBayes classifier to those vectors and classify each user into one of the three groups: SUPPORT, REFUTE, and JOKE. Through various parameter settings, we achieve the best result with 80% accuracy using 10 fold cross validation and the dimension of the vector is 200. The classification result agrees with the manually classified data based on the two human assessors and further supports our hypothesis about the intrinsic content characteristics of each user group.

VI. Discussion

The first research question asked if we could determine if there is a specific group of users that is greatly interested in discussing and spreading rumours based on user activities in Reddit. Our visualization tool allowed us to identify these users, and manual analysis of users in this highly-engaged category revealed persistent interaction with the rumour over 10 years for 163 users. As noted in our threats to validity section, it is possible to disguise high levels of interest in a particular rumour, but our approach is effective in many cases.

The second question asked if we could assess user activities and assign types to users based on whether they support, refute, or joke about the rumour. Our manual assessment of all user comments revealed that users did engage with the rumour in a consistent pattern. These users can be categorized into three groups: (1) "Generally support a false rumour", (2) "Generally refute a false rumour", and (3) "Generally joke about a false rumour". We further examined how users in these categories interacted, and found that joking users were the most active. Using both social network analysis and content analysis provided us with some interesting results. Users in the JOKE category seemed to be the most active group that interacted with the rumour among themselves and with users in other categories. The content of user comments in REFUTE was explanatory and fact-driven, while the content of user comments in SUPPORT lacked details and evidence. Finally, applying text mining techniques allows us to identify those users automatically.

The third question asked if visualizing the spread of a rumour could provide better insight into how users interact with rumours. The assessment of this is subjective and qualitative, but we certainly found that a visual depiction of how rumour-spreading users interacted with submissions, with comments, and with other users was helpful in tracing and understanding the spread of the rumour. The longitudinal analysis showed a persistent rumour, and identified submissions on the topic that deviated from others based on an automated semantic assessment. The visualization tool allows both a high-level view and a detailed breakdown; some visualizations are presented in this paper and are certainly helpful in drawing hypotheses that could not be driven without visual observation of raw data.

Threats to Validity

Our evaluation is based on a case study and our own observations about our method, both threats to external validity. Our results will require further validation before we can confidently assert that they apply generally.

Redditors can, and do, change their usernames, create new identities, or update/delete their own comments or submissions. Our counts of users engaging with this rumour is therefore a floor, not a ceiling or a precise measurement.

In this case study, we focus on a false rumour that was debunked by Snopes. We would also like to investigate if the proposed approach is still valid for rumours that are only partly false.

VII. Conclusions

In this paper, we presented a study about how users interact with rumours in Reddit. The results have shown that a specific group of users actively interacted with the chosen rumour. These users are categorized into three groups: (1) “Generally support a false rumour”, (2) “generally refute a false rumour”, and (3) “Generally joke about a false rumour”.

The use of social network analysis, content analysis, visualizations, sentiment analysis, and text classifications validate and support the proposed approach. Users in the category “Generally joke about a false rumour” seemed to be the most active group that interacted with the rumour among themselves and with users in other categories. The content of user comments in “Generally refute a false rumour” was explanatory and fact-driven, while the content of user comments in “Generally support a false rumour” lacked details and evidence. Finally, those users are grouped in one of the three categories automatically using text classification.

We illustrate our general approach using data from Reddit. This approach is also suitable for other OSNs (like Flickr or Twitter); however, OSNs do not always exhibit the same user behavior, so the specific results of our analysis are not necessarily true of other OSNs. Additional studies will be required to assess user behavior on other OSNs.

Acknowledgment

The research was funded in part by the Canada Foundation for Innovation, CNPq and FAPESP, Brazil, Natural Sciences and Engineering Research Council of Canada, and International Development Research Centre, Ottawa, Canada.

References

[1] R. L. Rosnow, “Inside rumor: A personal journey.” *American Psychologist*, vol. 46, no. 5, p. 484, 1991.

[2] H. Kwak, C. Lee, H. Park, and S. Moon, “What is twitter, a social network or a news media?” in *Proceedings of the 19th International Conference on World Wide Web*. ACM, 2010, pp. 591–600.

[3] E. Morozov. (2009) Swine flu: Twitter’s power to misinform @ONLINE. Available: http://foreignpolicy.com/posts/2009/04/25/swine_flu_twitters_power_to_misinform [Accessed: April 15, 2015].

[4] C. Chew and G. Eysenbach, “Pandemics in the age of twitter: content analysis of tweets during the 2009 h1n1 outbreak,” *PloS One*, vol. 5, no. 11, p. e14118, 2010.

[5] B. Doerr, M. Fouz, and T. Friedrich, “Why rumors spread so quickly in social networks,” *Communications of the ACM*, vol. 55, no. 6, pp. 70–75, 2012.

[6] M. Nekovee, Y. Moreno, G. Bianconi, and M. Marsili, “Theory of rumour spreading in complex social networks,” *Physica A: Statistical Mechanics and its Applications*, vol. 374, no. 1, pp. 457–470, 2007.

[7] E. Seo, P. Mohapatra, and T. Abdelzaher, “Identifying rumors and their sources in social networks,” in *SPIE Defense, Security, and Sensing*. International Society for Optics and Photonics, 2012, pp. 83 891I–83 891I.

[8] D. Shah and T. Zaman, “Rumors in a network: who’s the culprit?” *Information Theory, IEEE Transactions*, vol. 57, no. 8, pp. 5163–5181, 2011.

[9] P. Bordia, N. DiFonzo *et al.*, “Psychological motivations in rumor spread,” *Rumor Mills: The Social Impact of Rumor and Legend*, pp. 87–101, 2005.

[10] L. Lü and T. Zhou, “Link prediction in complex networks: A survey,” *Physica A: Statistical Mechanics and its Applications*, vol. 390, no. 6, pp. 1150–1170, 2011.

[11] T. Yang, R. Jin, Y. Chi, and S. Zhu, “Combining link and content for community detection: a discriminative approach,” in *Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, 2009, pp. 927–936.

[12] S. Lewandowsky, U. K. Ecker, C. M. Seifert, N. Schwarz, and J. Cook, “Misinformation and its correction continued influence and successful debiasing,” *Psychological Science in the Public Interest*, vol. 13, no. 3, pp. 106–131, 2012.

[13] T. Falkowski, J. Bartelheimer, and M. Spiliopoulou, “Mining and visualizing the evolution of subgroups in social networks,” in *Proceedings of the 2006 IEEE/WIC/ACM International Conference on Web Intelligence*. IEEE Computer Society, 2006, pp. 52–58.

[14] D. Liben-Nowell and J. Kleinberg, “The link-prediction problem for social networks,” *Journal of the American Society for Information Science and Technology*, vol. 58, no. 7, pp. 1019–1031, 2007.

[15] M. J. Brzozowski, T. Hogg, and G. Szabo, “Friends and foes: ideological social networking,” in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 2008, pp. 817–820.

[16] C. Buntain and J. Golbeck, “Identifying social roles in reddit using network structure,” in *Proceedings of the Companion Publication of the 23rd International Conference on World Wide Web Companion*. International World Wide Web Conferences Steering Committee, 2014, pp. 615–620.

[17] A. Dang, R. Makki, A. Moh’d, A. Islam, V. Keselj, and E. E. Milios, “Real time filtering of tweets using wikipedia concepts and google tri-gram semantic relatedness.” in *TREC*, 2015.

[18] L. Festinger, D. Cartwright, K. Barber, J. Fleischl, J. Gottsdanker, A. Keysen, and G. Leavitt, “A study of a rumor: its origin and spread.” *Human Relations*, 1948.

[19] K. Marett and K. Joshi, “The decision to share information and rumors: Examining the role of motivation in an online discussion forum,” *Communications of the Association for Information Systems*, vol. 24, no. 1, p. 4, 2009.

[20] R. M. Ryan and E. L. Deci, “Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being.” *American Psychologist*, vol. 55, no. 1, p. 68, 2000.

[21] V. Qazvinian, E. Rosengren, D. R. Radev, and Q. Mei, “Rumor has it: Identifying misinformation in microblogs,” in *Proceedings of the Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics, 2011, pp. 1589–1599.

[22] J. Ratkiewicz, M. Conover, M. Meiss, B. Gonçalves, S. Patil, A. Flammini, and F. Menczer, “Truthy: mapping the spread of astroturf in microblog streams,” in *Proceedings of the 20th International Conference Companion on World Wide Web*. ACM, 2011, pp. 249–252.

[23] A. Dan and J. Richards. (2011) Behind the rumours: how we build our twitter riots interactive @ONLINE. Available: http://www.guardian.co.uk/news/datablog/2011/dec/08/twitter-riots-interactive [Accessed: April 15, 2015].

[24] A. Dang, A. Moh’d, E. Milios, and R. Minghim, “What is in a rumour: Combined visual analysis of rumour flow and user activity,” in *Proceedings of the 33rd Computer Graphics International*. ACM, 2016, pp. 17–20.

[25] A. Dang, A. Moh’d, A. Gruzd, E. Milios, and R. Minghim, “A visual framework for clustering memes in social media,” in *Proceedings of the 2015 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2015*. ACM, 2015, pp. 713–720.

[26] E. Otte and R. Rousseau, “Social network analysis: a powerful strategy, also for the information sciences,” *Journal of Information Science*, vol. 28, no. 6, pp. 441–453, 2002.

[27] J. Zhao, N. Cao, Z. Wen, Y. Song, Y.-R. Lin, and C. M. Collins, “# fluxflow: Visual analysis of anomalous information spreading on social media,” *Visualization and Computer Graphics, IEEE Transactions on*, vol. 20, no. 12, pp. 1773–1782, 2014.

[28] W. Wang, H. Wang, G. Dai, and H. Wang, “Visualization of large hierarchical data by circle packing,” in *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*. ACM, 2006, pp. 517–520.

[29] A. Friggeri, L. A. Adamic, D. Eckles, and J. Cheng, “Rumor cascades.” in *ICWSM*, 2014.

[30] L. Dang-Xuan and S. Stieglitz, “Impact and diffusion of sentiment in political communication-an empirical analysis of political weblogs.” in *ICWSM*, 2012.

[31] W. Zhang and S. Skiena, “Trading strategies to exploit blog and news sentiment.” in *ICWSM*, 2010.