Research Article
Mining and Exploring the Structure of Academic Tweets

Norah Farooqi, AlaaAlmelibari and TahaniAlsubait
Umm Al-Qura University, Makkah, Saudi Arabia

Abstract: This paper aims to study the structure of tweets in academic accounts to ascertain the most effective form(s) of tweets. It determines the attributes which are usually used and evaluates the impacts on tweets' interactivities. Since the use of the micro blogging platform Twitter has spread globally and developed rapidly in the last few years. Users engage it in all aspects of their lives including education, economy and healthcare. Thus understanding data analytics in Twitter improves related environments, providing fruitful feedback, developing procedures and enhancing information management in organizations. Mining data types in tweets and following their feedback leads to clarifying the importance of each attribute. The reported findings present the recommended structure of tweets in academic environments due to their impact on users. This research offers a helpful guide to develop scholarly Twitter accounts and to support them to be more interactive.

Keywords: Academic accounts, mining attributes, social media, twitter, twitter analytics

INTRODUCTION

Social media affects our daily life markedly in many ways. Different tools with various features have been developed for social networking. Twitter is a micro blogging platform that was introduced in 2006 for sharing short messages. It has expanded rapidly during the last ten years to handle millions of accounts and tweets. Currently, it is considered as one of the primary methods for communication and advertisement in the majority of companies around the world. It can be used formally or informally in different ways according to the requirements and specifications of a particular application. The content of tweets is a rich source of valuable information that invites investigation and analysis.

It is essential to take note of the wide research gap in the field of Twitter analytics. The literature suffers from a limitation of scholarly sources discussing how Twitter analytics can be used in different fields. Besides, there are insufficient studies about the structure of tweets to determine the used-attributes. Therefore, in light of this, this proposed research focuses on studying the structure of tweets in the educational sector. It measures the relationships between different data types such as Text, Image, Video, Tags and URLs within a tweet and the user's reactions to those tweets. The proposed research will also seek to understand how the structure of tweets influences the number of followers in a particular account.

LITERATURE REVIEW

Bhulai*et al*. (2012) discuss a study carried out by the Government Engineering College, India, whose main aim was to understand how engineering departments could benefit from Twitter mining. According to the authors, following the conversations on Twitter can shed light on the opinions, educational experiences, concerns and feelings related to the process of learning. Through its data mining, the department was able to uncover that lack of social engagement, heavy study load, as well as sleep deprivation were some of the main issues that engineering students faced. This study presented an opportunity for the department to amend the regulations that will reduce the study load for their students, improve sleep time and ensure social engagement (Bhulai*et al*., 2012). For institutions that are facing different challenges but are yet to understand their root courses, they are encouraged to turn to Twitter data mining, which will make it easier to understand what affects such students and such analyses will be certainly cost-effective.

Carpenter *et al.* (2014) acknowledge that communication has long been an integral part of educational leadership. More importantly, they note that such communication is not confined to school personnel but extends to the wider community. From general interactions in any community, communication is more effective and useful if it entails feedback that, in turn, determines the next course of action (Carpenter

Corresponding Author: Norah Farooqi, Umm Al-Qura University, Makkah, Saudi Arabia

et al., 2014). From this, the most apparent benefit of Twitter in education stems from the concept of communication and generation of feedback. Agrawal *et al*. (2012) in her publication categorized feedback in education into:

- Feedback to the instructor from the students
- Feedback to the students from the instructor

The first gives students the opportunity to guide the instructor towards teaching in a way that they find most understandable, while the second is for the students' self-improvement. Typically, schools use Student Response Systems (SRS) in classrooms for feedback, given to the instructors by the students, through devices such as smart phones and clickers (Agrawal *et al*., 2012). However, Leong *et al*.(2012) argue that SRSs are characteristically not as effective in the provision of feedback on matters that could have gone wrong as are social media platforms. More specifically, the researchers point out Twitter as one of the most effective social media platforms that students can use to provide concise yet precise feedback. Zeng *et al*. (2014) described sentiment analysis as a field that attempts to make sense out of opinions and textual feedback. Moreover, Denker (2013) pointed out that views can either be positive or negative and are associated with a wide range of emotions. The resultant negative emotions include feeling irritated, while positive ones are expressed by the students when they tweet louder writing styles, for example, bold writing (Denker, 2013). However, Carpenter *et al*. (2014) point out that there is also neutral feedback in which students display neither positive nor negative emotions.

In his publication, Taboada *et al*.(2011) argues that Educational Data Mining (EDM) is a critical area of research that is instrumental in improving education because of its ability to monitor and understand students' learning and performance. In the agreement, Porterfield and Carnes (2012) add that collecting students' feedback at the end of the semester after already taking the course does not add value. Rather, they argue that feedback is more helpful to the current students when it is collected in real time and, equally importantly, addressed in real time. Earlier, studies by Cummings and Hsu (2007) had shown that when students' feedback is analyzed using techniques of sentiment analysis, it is possible to identify not only their negative or positive feelings but even refined emotions that they have with regards to the current teaching can also be illustrated. Vohra and Teraiya (2013) further developed the findings by Taboada *et al*.(2011) and demonstrated that by addressing the issues revealed by EDM, students could be advised according to their specific needs. Equally importantly, the material can be added or removed to the unit

following the students' opinion of the course and their level of comprehension.

Multiple studies have been conducted to understand the benefits of Twitter in education. Wang (2013) began by exploring the significance of the concept of student feedback and the platform through which it is collected. In agreement with earlier findings by Agrawal *et al*. (2012), he found that student feedback is critical because it helps not only instructors but also curriculum developers to gain deeper insight into the learning behavior of students. While clickers help students concentrate on the material for longer, they are relatively costly and potentially distract them from learning. Alternatively, mobile phones are popular among students and instrumental in sending feedback. However, students also tend to use their phones unhelpfully in class, consequently overflowing the system with inappropriate content and causing distractions. With regards to Short Message Service (SMS), studies by Denker (2013) showed that the cost of the service limited the amount and quality of feedback that students offered. However, Wang (2013) found that Twitter is advantageous for several reasons that stem from the core the fact that it is a familiar tool among more than 95.6% of the student population and does not require training in using it.

Carpenter *et al*. (2014) also argue that Twitter benefits education because it provides free feedback and students can open it on their mobile phones via the wireless networks of their institutions. On the one hand, the limited number of characters may be viewed by critics of Twitter as a disadvantage in the sense that feedback is less detailed. However, Leong *et al*. (2012) argue that the limited characters compel the students to focus on key issues and create the most meaningful sentence possible. Then, using sentiment analysis that facilitates opinion mining, appraisal extraction and subjectivity analysis for natural language processing, the instructors can identify and retrieve specific information from the tweets. As already argued by Carpenter *et al*. (2014), feedback in education will enable instructors to measure not only their students' but also the public's opinions, perceptions and moods, which are crucial in making decisions. It is also imperative to note that the participation of educational leaders in such communication will enable them to exercise a degree of message control. With specific regards to public schools in London and Cardiff, Porterfield and Carnes (2012), the analysis showed that school leaders and schools that do not participate in Twitter miss out on what their communities think and say about them. It follows, therefore, that the consequence of not knowing how an institution is perceived is that no corrective or preventative actions are taken.

Earlier, Cummings and Hsu (2007) had argued that leaders of educational institutions who do not

participate in Twitter are comparable to the situation a decade ago in which school leaders were never available for meetings with parents or other stakeholders. Currently, educational leaders who do not participate in Twitter (or any other social media platform) do not respond to requests by parents and do not maintain beneficial relationships with leaders. Once leaders and their institutions are viewed as nonresponsive and lacking social sensitivity, the risk of becoming the subject of intensive media scrutiny and losing community support is real (Cummings and Hsu, 2007). Vohra and Teraiya (2013) argue that social networks have essentially become a culture with which humanity identifies. Without being left behind, academic organizations, institutions and societies have increasingly become more connected. With regards to the categorizations of feedback by Agrawal *et al.* (2012), feedback to the instructor from the students appears to be the origin of the benefits of Twitter to the education sector. However, while students are the primary beneficiaries of the education system, it would not be accurate to limit the impacts of communication only to that which is generated by the students (Wang, 2013).

In his explanation of what education institutions should consider when taking Twitter beyond the boundaries of the institutions themselves, Wang (2013) argues that it is imperative to define what the tweets are intended to achieve. In that way, the benefits can be quantified. Equally importantly, the institution must consider who they want to connect with and what value they can deliver to the followers they wish to attract. Evidently, the values an institution delivers will determine not only the type but also the number of followers they attract. While student feedback will directly influence instructional methods, feedback and contributions from external followers will impact the general management of the institution. However, Mathioudakis and Koudas (2010) argue that the general supervision of the institution cannot be isolated from the instructional methods

With regards to the importance of analyzing Twitter data, Taboada*et al*. (2011) argue that it is for the facilitation of making the most relevant and effective decisions. As already shown in the previous works, sentiment analysis entails evaluating moods or emotions with regards to the data under analysis (Zeng *et al*., 2014). Therefore, sentiment analysis, using natural language processing to analyze the use of words, their order and combination, analyses the sentiments expressed by Twitter users to measure the general sentiment. The most common sentiment analysis techniques pointed out by Carpenter *et al*. (2014) are Naïve Bayes (NB), Support Vector Machines (SVM) and Max Entropy (MaxEnt). NB gives accurate results when it is implemented on blog posts and reviews, while MaxEnt provides mixed

(accurate and inaccurate) results. SVM provides lower degrees of accuracy when implemented on its own on reviews, but produces higher accuracy when used in combination with NB and MaxEnt. However, the essence is that social media analytics are crucial in facilitating the detection and solving of problems.

Studies by Leong *et al*. (2012) show that wealthier and larger school districts are relatively more prominent on Twitter and the substance of their communications and feedback are generally positive. On the other hand, they also show that for smaller school districts, there may be smaller numbers of followers, but they typically communicate not only repeatedly but also in concentration. More specifically, Vohra and Teraiya (2013) point out that while the topics the followers' addresses are generally broad, they tend to gravitate towards people, performance, operations, policy, safety, controversy and politics. Carpenter *et al*. (2014) agree but specifically note that people form the most prevalent theme while performance is the worst. However, it is more important to note that despite what the followers tweet, they fundamentally express their opinions, which can provide critical data that the educational institutions can use to improve themselves.

Narrowing the focus down to how Twitter can be used in education by students and instructors, Porterfield and Carnes (2012) pointed out several benefits. They argue that Twitter can be used to send assignments; for student collaboration; to subscribe to relevant hashtags; as a research tool; and as a tool for surveying students. In their explanation, they add that both students and instructors can tweet assignment details quickly, conveniently and easily via Twitter in education. Links to relevant online sources can also be included, whereby every person with a Twitter account will receive tweets to which they can respond, as well as share or forward. Porterfield and Carnes (2012) also add that email is rapidly being challenged by Twitter as a collaboration tool in schools. Since instructors can motivate their students to work on collaborative assignments using Twitter, they also have the ability to intervene at any point and offer positive advice, share relevant links and help students to identify appropriate resources. However, Porterfield and Carnes (2012) add that it is imperative for instructors to discuss with their students the issues they would be interested in following on Twitter. The essence of such discussion is to involve everyone in deciding the hashtags they should all follow. Consequently, by receiving current regular news updates, the entire group will always be in tandem when discussing the topics in class.

The above sections have sought to understand how the structure of tweets influences the number of followers in a certain account by undertaking a review of past literature on the same topic. It has discussed different network measures according to previous findings from other researchers and has addressed who

is considered as important when it comes to generating data from Twitter. Importance is judged in line with the users who generate the most tweets, the users who are the most influential and the users who control the flow of information (Smith and Barash, 2008). Different analytic methods are applied in achieving this including Eigenvector centrality, betweenness centrality and degree centrality. The section had also discussed how this data could benefit different sectors and institutions by reviewing the findings of past literature. In general, institutions can use such information to understand different market trends and the demands of their clients in order to improve their effectiveness and success (Kang *et al.*, 2011).

With regards to the education sector, the above section has shown that the most notable benefit is the concept of communication and generation of feedback. Feedback can travel from instructors to students for their (students') self-development or from the students to the instructors to guide them on how to deliver understandable instruction (Plachouras and Stavrakas, 2012). It has also been shown that educational data mining has the potential to improve education through its ability to monitor and understand both the learning and performance of students. The essence of combining the analysis of the feedback and the monitoring of students is that school leaders will have critical information on which to base decisions designed to benefit the education sector. In contrast, schools and school leaders who do not participate in social media compromise the platforms on which they can receive information on how their institutions are perceived by the wider community.

MATERIALS AND METHODS

This section focuses on the proposed study approach and justifies the methods used to achieve the experimental results.

Methodology: As previously stated, this study is mainly concerned with understanding the tweet message structure, analyzing data types and ascertaining the most useful forms of tweets. Tweet messages are categorized as original, retweeted and reply. An original tweet message has its data and information that are published by the main account. In contrast, a retweeted message repeats data that have been published by other accounts. A reply message is written by the main account to interact with other accounts according to the situation. Every tweet contains one or more attributes that build up its structure and receive one or more interactions as feedback. Attributes include Text, Hyperlinks, Images, Videos, Emojis, Mentions, Hashtags and Tags in photos. Interactions for tweets include retweet, like and reply. The general structure of a tweet message is illustrated in Fig. 1.
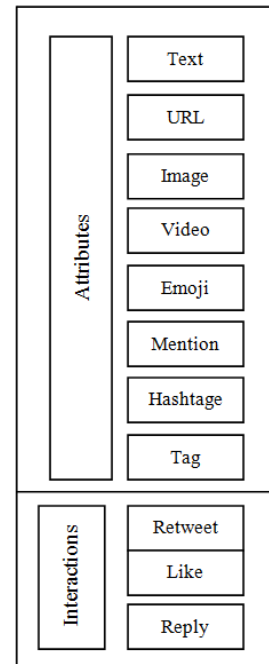

Fig. 1: The basic structure of a tweet

This study adopts a quantitative methodology to determine the proper data structure for tweets on Twitter in the educational sector. It was conducted in five main stages:

- **Analyzing accounts:** Tweets in each single Twitter account were evaluated separately and these results were processed to determine the common attributes that were used.
- **Measuring attributes:** Each attribute in the messages' structure was measured and tested for all selected Twitter accounts to determine the importance of the attributes.
- **Aggregating attributes:** The volume of each attribute was calculated over all selected Twitter accounts to compare usability.
- **Comparing interactions:** Interactions for evaluated tweets in each account were collected to ascertain the general feedback.
- **Evaluating correlations:** Each attribute was related to the total interactions for each account to calculate the correlation value between data types and interactions.

Experimental work: This section explains the experimental setup for the proposed approach. The main goal of these experiments is to evaluate Twitter accounts in the education sector and analyze the messages' structure. All experiments were performed on a PC with 1.70 GHz Intel® Core™ i5 CPU, 4 GB of main memory and Windows 8 operating system. Data sets were collected using the Twitter API that depends on Tweepy and Python.

Table 1: Basic information about the top ten universities Twitter accounts in the UK

Abbreviation	University name	Twitter account	Joined date	Followers	Tweets	Likes	Media count	Collection date
U1	University of Oxford	@UniofOxford	June 2009	= 357K	9071	3035	3903	7-4-2017 To 17-4-2017
U2	University of Cambridge	@Cambridge_uni	April 2009	= 323K	7276	1086	917	6-4-2017 To 16-4-2017
U3	Imperial College of London	@imperialcollege	July 2008	= 59K	7733	2428	3210	2-4-2017 To 17-4-2017
U4	University College London	@ucl	June 2011	=31K	1803	286	797	6-3-2017 To 12-4-2017
U5	London School of Economics & Political Sciences	@LSEnews	June 2010	=45K	=10.4 K	1028	3844	6-4-2017 To 17-4-2017
U6	University Of Edinburgh	@EdinburghUni	March 2009	=76,6K	5320	506	373	19-3-2017 To 17-4-2017
U7	King's College London	@KingsCollegeLon	March 2009	=61,6K	5274	601	684	15-3-2017 To 15-4-2017
U8	University of Manchester	@OfficialUoM	November 2013	=27,8K	8652	556	5240	4-5-2017 To 16-4-2017
U9	Bristol University	@BristolUni	March 2009	=66K	=10,5 K	1222	970	27-3-2017 To 16-4-2017
U10	University Of Warwick	@warwickuni	August 2008	=64,8K	=18K	2276	2034	6-4-2017 To 13-4-2017

The empirical dataset contains Twitter accounts for the best ten UK universities according to the Times Higher Education ranking. The basic information about these Twitter accounts is illustrated in Table 1; this was gathered on 30-4-2017. This information focuses mainly on some attribute measurements that determine the interactive status of each account.

The join measurement shows the beginning of university activities in the microblog of the social media platform Twitter. Both the Imperial College of London and the University of Warwick created their accounts in the middle of 2008 and subsequently, other universities started to activate their accounts between 2009 and 2011. The University of Manchester was the last to open its account, which occurred in November 2013. The second recorded measurement in Table 1 is the number of followers to understand the engaged size of society for each account. The maximum number of followers among these selected top ten UK universities is for the University of Oxford, which reaches 357K followers. The minimum number of followers is for the University of Manchester, which has approximately 27.8K followers and the reason for this may be the late start to its Twitter presence, as mentioned before. The third measurement of interaction is the volume of tweets for each account. It seems that the University of Warwick, which started very early, is the most interactive university from the list since its tweets number more than 18K.

On the other hand, University College London has the fewest number of tweets, around 1,803. The volume of tweets for the remaining universities varies from one account to another. The measurement of 'likes' is also selected because it reflects the interaction in each account. The University of Oxford has the largest volume of likes and the University College London has

the smallest volume of likes. In general, in comparing the number of tweets with the number of likes, the volume of tweets is always larger than the volume of likes for all accounts. The rest of this study will refer to universities' accounts using abbreviations (e.g., 'U1' to represent the University of Oxford) for simplification.

The sample of evaluated tweets is taken to be 1,000 tweets, which is divided into 100 tweets for each university account. The tweets included the last 100 tweets until 17-4-2017 and the period was approximately one month. The collected data contains three main parts for each tweet. The first part includes the basic information about the selected tweets that shows the issued date and its type (original, retweet, or reply). The second part reflects the recorded interaction for specific tweet messages by measuring the number of retweets, likes and replies. The third part analyses the tweets' message content and categorizes its structure into Text, Hyperlink, Image, Video, Emoji, Mention, Hashtag and who is in the photo.

RESULTS AND DISCUSSION

This study focuses on mining the tweet structure for educational uses and analyzing attributes. The results of applying this approach are described in this section. Experiments were conducted in many stages as described previously. The outcomes of each stage are explained sequentially.

Analyzing accounts: The results of analyzing the tweet structure in each single Twitter account are illustrated in Fig. 2. Each university used attributes in different ways to build various tweet structures. The majority of universities utilize all attributes, with various
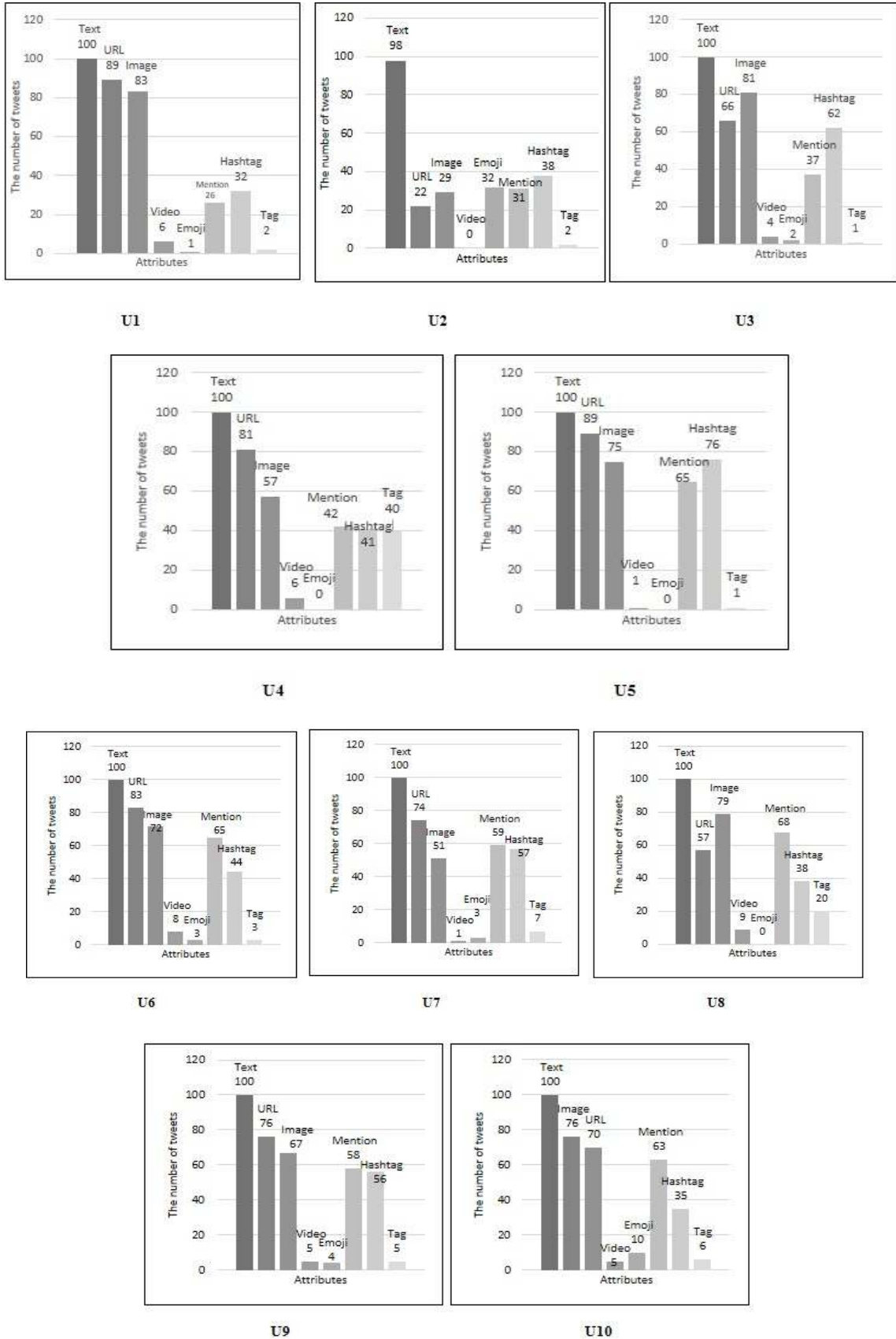
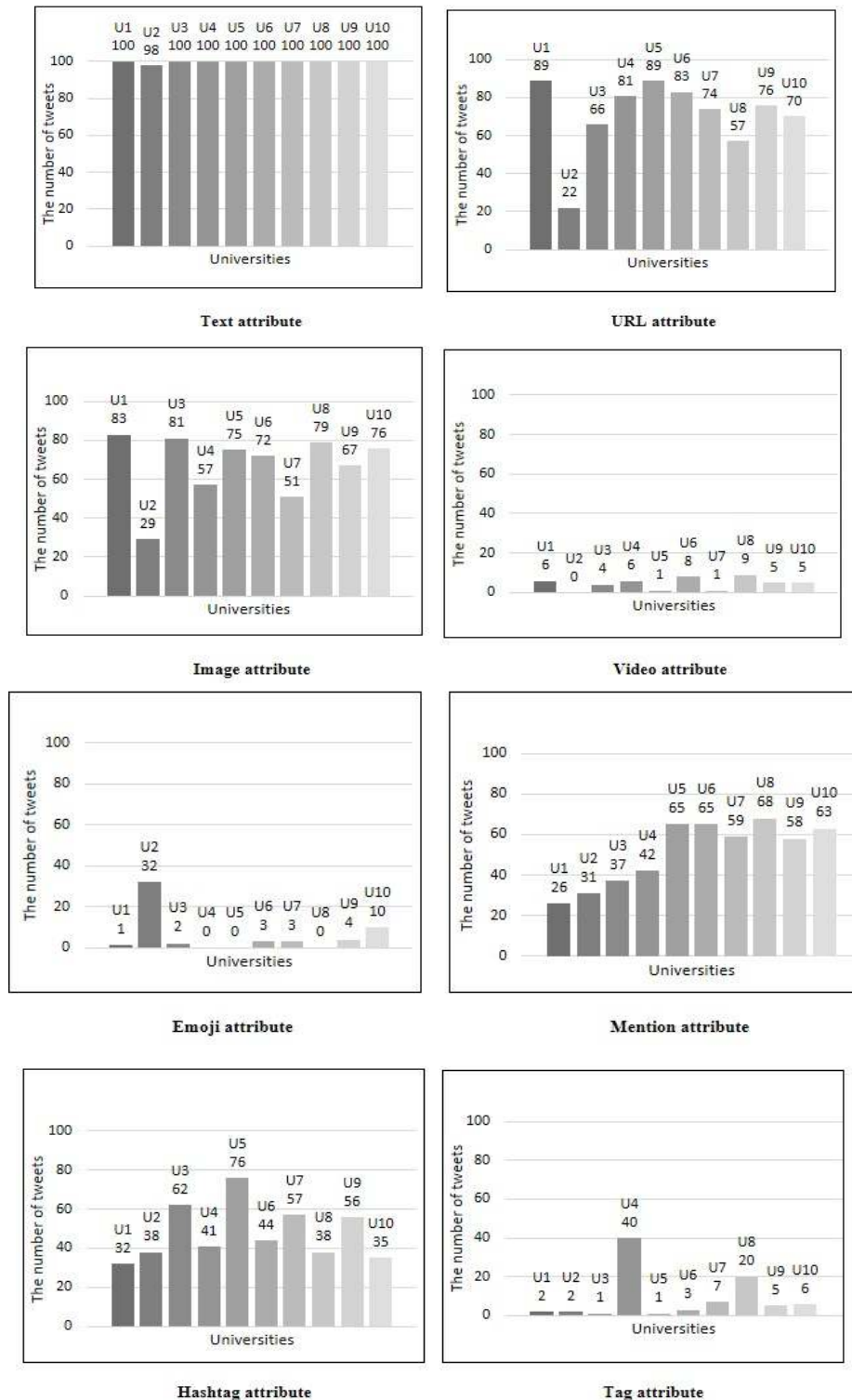Fig. 2: The results of analysis of the tweet structure

Fig. 3: The results for attributes

percentages. However, some universities did not use some attributes at all in the collected datasets. For example, U4, U5 and U8 did not use the Emoji attribute. The ranking of using attributes is similar in U1, U3, U4, U5, U6, U8, U9 and U10. In contrast, in U2 and U7, the order for attributes is quite different. By comparing data statistics between the ten universities' accounts, the most commonly used attribute is Text, followed by URL and Image. Tracking tweets of the educational sector on the Twitter platform gives a general understanding of tweets' structure commonly used in educational messages.

Measuring attributes: This stage explains the presented results from a different perspective in Fig. 3. It displays each attribute in all selected Twitter accounts to determine to what extent these attributes are important. The results for Text shows that it is very permanently used in all universities. It also seems that both URL and Image are highly used in the majority of universities. The Video is occasionally used and did not exceed ten uses in any account. Emoji is rarely used in all universities except in U2. Both Mention and Hashtag are utilized in the Moderate range. Both U4 and U8 frequently used the Tag attribute to explain who is pictured in photos. The results show the importance of each attribute based on tweets' structural analysis.

Aggregating attributes: The results for total uses of attributes overall selected Twitter accounts are explained in Fig. 4. The Text attribute is used 998 times out of 1,000 tweets, equal to 99.8%. The URL attribute is used in 707 tweets of 1,000 tweets (70.7%). The Image attribute is utilized in 670 tweets out of 1,000 (67%). The Mention attribute is used in around half of the tweets (51%). The Hashtag attribute is used in 479 out of 1,000 tweets, which is 47.9%. The Tag-Users attribute is used 138 times (13.8%). Finally, the Emoji attribute is used 55 times (5.5%) and the Video attribute is used in only 45 tweets out of 1,000 (4.5%).

According to these percentages, we can classify the usability of attributes as fundamental, secondary and additional. The fundamental attributes in the tweet structure are associated with percentages of more than 70% and include Text and URL. The secondary attributes have values of more than 50% and less than 70% and these include Image and Mention. The additional attributes have values of less than 50% and include Hashtag, Tag Users, Emoji and Video. These results reflect the use of different attributes to measure users' familiarity with them in all selected accounts.

Comparing interactions: This stage focuses on understanding the relationship between interactions and tweets for general feedback. The results compare collections of interactions for selected tweets in all universities. The volume of interactions includes the number of retweets, likes and replays.

Table 2 shows that the most interactive account is U1, with 6,765 interactions for 100 tweets. The least interactive account is U8, with only 886 interactions for 100 tweets; this may be due to the limitations in the number of followers. However, relating the volume of interactions to the number of followers is not sufficient. For example, U7 has 61.6K followers and shows 4,009 interactions, but U2 has 323K followers and reports fewer interactions, approximately 2,974. The remaining universities' accounts interact with 100 tweets in various formats. In general, it appears that the highest number of interactions for all universities involves likes followed by retweets and it is seen that replay interaction is very limited. The full corresponding interactions show to what extent each account is active.

Table 2: Interactions for 100 tweets

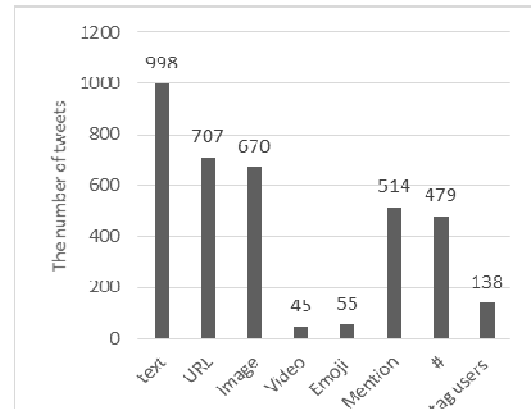| Name	Retweets	Likes	Replay	Total
U1	2218	4439	108	6765
U2	716	2042	216	2974
U3	656	870	13	1539
U4	1508	1713	79	3300
U5	416	495	18	929
U6	1750	2254	63	4067
U7	1455	2449	105	4009
U8	305	561	20	886
U9	781	1143	23	1947
U10	429	510	15	954

Fig. 4: Total uses of attributes overall selected accounts

Evaluating correlations: This section studies the relationship between interactions and the structure of tweets. The value of the correlation coefficient is between −1 and +1 and is calculated for each attribute from the total interactions in each tweet. A minus value indicates a negative correlation and a plus value shows a positive correlation. Every specific tweet is evaluated separately and each attribute is assigned a value of 1 if it exists or 0 if it does not exist. The attribute value indicates the total number of interactions for each related tweet. The results for the correlation coefficient are explained inTable 3. Some attributes do not relate sufficiently to the number of interactions due to the limited diversity present and the small difference between 1 and 0. In these cases, it is referred to as 'not applicable' with the abbreviation N/A.

For example, the Text attribute is always present in all tweets and the diversity value for its existence is minimal, so it cannot be used as an influential factor for increasing or decreasing the total interactions. Although the URL attribute is highly used by the majority of universities, it has negative correlations with interactions. Seven out of 10 universities have positive correlations between utilizing the Image attribute and showing an increased number of interactions for tweets. All universities use the Video attribute minimally, but it has a high impact regarding increasing the number of interactions. Five universities out of 10 have very positive correlations between using Video and the numberof interactions.In contrast, it appears that the

Table 3: The correlation between total interactions and each attribute

Name	Text	URL	Image	Video	Emoji	Mention	Hashtag	Tag
U1	N/A	-0.333959681	-0.01951	0.20464	0.077568	-0.05876	0.071045	0.025947
U2	0.054963	0.071061745	0.273604	N/A	-0.20469	0.12093	0.174816	0.017406
U3	N/A	-0.370319419	0.202841	0.107514	-0.04735	-0.14619	0.318608	0.014761
U4	N/A	-0.125772666	0.182206	0.31228	N/A	0.038025	0.352043	0.136554
U5	N/A	-0.275991018	-0.21661	-0.01765	N/A	0.003525	-0.09252	0.025263
U6	N/A	-0.001588328	-0.07664	0.294327	-0.04392	-0.15074	0.160475	-0.03493
U7	N/A	-0.374489054	0.111356	0.206962	0.008291	0.091041	0.084448	0.023162
U8	N/A	-0.019416029	0.332235	-0.15333	N/A	0.054539	-0.07447	0.234554
U9	N/A	0.008933277	0.172454	-0.01391	-0.01347	0.00419	0.04181	0.203067
U10	N/A	-0.184666231	0.129695	-0.12822	0.05883	0.159296	0.076139	0.384727

Emoji attribute is not an important factor for increasing interactions for tweets. Mention and Hashtag have positive correlations, generally with various values. Nine out of 10 universities have positive correlations between the Tag Users in Photo attribute and an increasing number of interactions.

The results show the correlations' evaluation between data types and interactions. To conclude, some attributes that are highly used have negative correlations. In contrast, some other attributes that are not used frequently have positive correlations. Universities' accounts can increase interactions by using presentation attributes including Image and Video. It is highly recommended to apply reference-related attributes in tweets including Mention, Hashtag and Tag Users to make the accounts as interactive as possible. Also, it is recommended to reduce the use of the URL attribute due to its low effects regarding interactions.

The discussed results focused on tweet's structure in education while to the best of our knowledge the majority of the previous studies handled the tweet from social perspectives and analyzed the content sentimentally. Besides, there is a lack of scholarly sources discussing how Twitter analytics can be used in different fields such as education, health sector and economics. Thus, it is essential to consider the wide research gap in the field of Twitter analytics. In light of this, the research adds data that can be used to support twitter structural analysis highlighting how the information posted on Twitter can be used to improve different sectors.

CONCLUSION

Twitter has a significant impact on the daily lives of millions of users. Tweets analytics offers a multitude of ways to understand users' activities and to enhance the use of social media in the majority of aspects of living. This research focused on studying the structure of tweets in the academic sector, determined attributes and evaluated the related user interactions. The study was conducted in five main stages: Analyzing accounts, measuring attributes, aggregating attributes, comparing interactions and evaluating correlations.

The findings showed the most used attributes and how these affect users' interactions. The analysis of data demonstrated that the most commonly used attribute in the selected dataset is Text, followed by URL and Image. In contrast, the least commonly used attribute is Emoji, while Video was occasionally used and both Mention and Hashtag were in the Moderate range. Furthermore, the majority of university accounts tend to engage in 'like' more than 'retweet' and 'replay' interactions. The correlations between attributes and interactions showed that some attributes that are commonly used, such as URL, did not have a significant impact on increasing interactions. In contrast, the Video attribute, which is occasionally used, affects the number of user interactions. The recommended structure for tweets in the academic sector includes Text, Image and Video. To obtain highly interactive accounts and engage users, the accounts must always use reference-related attributes: Mention, Hashtag and Tag Users. This research will expand further to cover more comparative results for the academic sector in different countries and cultures. This approach of analyzing the structure of tweets will be applied in future studies for various additional sectors, including the health and financial sectors.

REFERENCES

Agrawal, R., S. Gollapudi, A. Kannan and K. Kenthapadi, 2012. Data mining for improving textbooks. ACM SIGKDD Explorations Newsletter, 13(02): 7-19.

Bhulai, S., P. Kampstra, L. Kooiman, G.Koole, M. Deurloo and B. Kok, 2012. Trend visualization on Twitter: What's hot and what's not? Proceeding of the1st International Conference on Data Analytics, pp: 43-48.

Carpenter, D.M., J.W. Robertson, M.E. Johnson and S. Blum, 2014. Social media analytics in education: What is it, how is it useful, and what does it tell us about how schools are discussed in social media? J. School Pub.Relat., 35: 7-43.

Cummings, R.G. and M. Hsu, 2007. The effects of student response systems on performance and satisfaction: An investigation in a tax accounting class. J. College Teach. Learn., 4(12): 21-26.

Denker, K.J., 2013. Student response systems and facilitating the large lecture basic communication course: Assessing engagement and learning. J. Commun. Teach., 27(1): 50-69.

Kang, U., D.H. Chau and C. Faloutsos, 2011. Mining large graphs: Algorithms, inference, and discoveries. Proceeding of the 27th International Conference on Data Engineering (ICDE'11), pp: 243-254.

Leong, C.K., Y.H. Lee and W.K.Mak, 2012. Mining sentiments in SMS texts for teaching evaluation. Exp. Syst.Appl., 39(3): 2584-2589.

Mathioudakis, M. and N. Koudas, 2010. Twitter monitor: Trend detection over the twitter stream.Proceeding of the 2010 ACM SIGMOD International Conference on Management of Data (ACM), pp: 1155-1158.

Plachouras, V. and Y. Stavrakas, 2012. Querying term associations and their temporal evolution in social data. Proceeding of the 1st International VLDB Workshop on Online Social Systems (WOSS 2012), Istanbul.

Porterfield, K. and M. Carnes, 2012. School communication in the age of Google. J. School Pub.Relat., 33(2): 115-130.

Smith, M. and V. Barash, 2008. Social SQL: Tools for exploring social databases. IEEE Data Eng., 31: 50-57.

Taboada, M., J. Brooke, M. Tofiloski, K. Voll and M.Stede,2011. Lexicon-based methods for sentiment analysis. Comput. Linguist., 37(2): 267-307.

Vohra, S. and J. Teraiya, 2013. Applications and challenges for sentiment analysis: A survey. Int. J. Eng. Res. Technol., 2(2): 1-5.

Wang, Y., 2013. Social media in schools: a treasure trove or hot potato?J. Cases Edu. Leadership, 16(1): 83-91.

Zeng, D., H. Chen, R. Lusch and S.H.Li, 2014. Social media analytics and intelligence.IEEE Intell. Syst., 25(6): 13-16.