Analyzing Public Discussions About #SaudiWomenCanDrive Using Network Science

ZUBAIDA JASTANIA1, RABEEH AYAZ ABBASI2, MUHAMMAD AHTISHAM ASLAM1, TARIQ JAMIL SAIFULLAH KHANZADA1,3, AND KHAWAJA MOYEEZULLAH GORI4

1Faculty of Computing and Information Technology, King Abdulaziz University, Jeddah 21589, Saudi Arabia
2Department of Computer Science, Quaid-i-Azam University, Islamabad 45320, Pakistan
3Department of Computer Systems Engineering, Mehran University of Engineering and Technology, Jamshoro 76062, Pakistan
4Department of Computer Science, National University of Modern Languages, Islamabad 44000, Pakistan

Corresponding author: Rabeeh Ayaz Abbasi (rabbasi@qau.edu.pk)

Abstract: Twitter has become a common place to share opinions about various issues. One such issue was the driving of women in the Kingdom of Saudi Arabia (KSA). Many researchers have addressed this issue using methods mostly from social sciences, however, these studies are on a smaller scale. In this paper, we take the approach of understanding the public discussions related to the topic #SaudiWomenCanDrive using methods from network science. Network science allows us to understand the influential people, topics, and communities related to the topic on a large scale. Around two million tweets posted by 630 thousand users are analyzed. Retweet, mention, co-mention, and hash-tag networks are extracted from these tweets. Analysis shows that most influential users include women rights activists like Loujain Hathloul and Manal Alsharif, as well as the political leadership like King Salman bin Abdulaziz and government departments like Ministry of Interior (MoI) and Traffic Department (Moroor). Most important topics identified through hashtags include the support for the movement like अल्लाह ने निर्देशित की महिलाओं का उद्योग (The King Supports Women’s Driving) as well as against it like أهل الشعب يرفض قيادة المرأة (Saudis Refuse Women’s Driving). Analyzing such discussions is important for gaining insights into public discussions for understanding public concerns and can be used by the government and non-government organizations to address them.

Index Terms: Network science, social network analysis, public discussions, social computing, social media analytics, Twitter, arabic.

I. INTRODUCTION

Twitter is one of the top leading microblogging social networking podiums. Often Twitter users use hashtags to categorize the same contents. Consequently, Twitter users, can use Twitter features of tweet, retweet, mention, co-mention, and hashtags to make the searching more comfortable, and provide an environment for further discussion. Comprehension of individual social behavior is essential to understand daily life processes. Network Science – evolved from graph theory – has been developed as a vital means for investigating social movements. The term “Social Network Analysis” (SNA) is often used interchangeably with the term “Network Science”. Analysis of Twitter data have several applications that have profound impact on human life [1], [2], including student and patient lives [3], understanding socio-economic aspects [4] and research communities [5]. Social networks, reflected by major human interaction, are deemed as the crucial basis of social behavior.

About 41% of the internet users in Saudi Arabia used Twitter in 2013.1 More recent statistics show that around 12.45 million users in Saudi Arabia use Twitter,2 that amounts for 36% of the entire population (including and excluding internet users). Therefore, analyzing the Kingdom of Saudi Arabia’s (KSA) Twitter landscape would help in gaining insights in user behavior on social media. The “Saudi Women Driving” dialogue was discussed for a vast span of time [6], [7]. For those who support the decision, they see that the decision is a right for women to serve themselves, gain independence, and to provide support to their families.

The associate editor coordinating the review of this manuscript and approving it for publication was Chunsheng Zhu.
In June 2018, Saudi Arabia removed the legal ban of women driving. Consequently, there were enormous consequences for transportation, low energy consumption, and labor competition between males and females [8]. The information flow on Twitter hides patterns of shared practice by Twitter users. Therefore, the information could be categorized into divisions, unified, fragmented, clustered, and in/out hubs [9]. Unlike predictive machine learning models [10] this research analyzes a trend of the topic without the need for any pre-training. Moreover, network science measures could reveal interesting patterns of users. For example, the density metric captures the interconnectivity between users, and the centrality metrics measure the importance of the Twitter user. Furthermore, the modularity can be used to cluster information based on community interests. Consequently, people are categorized into several categories, such as broadcast and support, and brand and clustered community networks [9].

Although there is an immense amount of literature in Twitter analysis, to the best of our knowledge a little work has analyzed a large-scale structure of the Arabic Twitter datasets related to the public discussions about allowing women to drive in Saudi Arabia [11]. Therefore, there is an immense need to investigate insights from the heated debate incurring through massive amounts of tweets. Some researchers have analyzed sentiments about the decision [11]. Authors collected 4,098 tweets and analyzed their sentiments. They found that around 67% of the tweets had a clear stance and majority of these tweets were in favor of the decision to allow women to drive. Although, the analysis done in [11] is based on a limited set of tweets, as compared to the number of tweets (two million) we have used for analysis, the results do overlap in the context that majority of the people support the decision, as discussed in Section IV(C). Other researchers have used methods commonly used in linguistics [12] to understand this discussion. The paper analyzes the discussions in a socio-cultural context and hence the results are not comparable to our research. In this research, we analyze and visualize results to know about the influential people, communities, important topics, and relationships among all these entities using methods from network science on a large scale.

The main objectives of our proposed approach are to find which users were more crucial in inevitable conversations, their opinion produced clusters or groups or not, and whether the more prevailing impact is Twitter reputation – such as, number of followers. The research further investigates user behavior analysis, like retweets and/or co-mentioned persons.

To analyze public discussions about #SaudiWomenCanDrive, we use methods from network science to analyze and comprehend human behavior about a public discussion. We analyze the relationship between users and tweets. As a result, the original data set that is extracted from Twitter about #SaudiWomanCanDrive and is converted into four sorts of networks: Mention-Network, Co-Mention-Network, HashTag-Network, and Retweet-Network. This paper analyzes co-mention and hashtag networks, whereas our earlier publication discusses the other networks. Each type of network reveals a piece of specific information about the top influential users, cohesive groups, and information disseminators.

II. RELATED WORK

Twitter analytical models can be categorized into three categories: machine learning approaches, custom-based use case models and metric-based approaches. The machine learning models are based on learning algorithms to predict, cluster and group Twitter users and topics. Twitter predictive models offer various ways for researchers to device and analyze reliability of public opinion [10]. Therefore, it is not only analyzing the current topics but moreover, anticipate the patterns and results of few acute real-world events such as gun violence and drug misuse. The emotional topics are not only capable of providing insights in community discussions, but also comprise of the machine learning algorithms that analyze Twitter content form linguistic and psychological perspectives [13]. Most machine learning models aim to identify and verify Twitter users [14]. This research analyzes discussions about a topic to gain insights into it. Twitter has been utilized for investigating opinions (sentiments) on numerous social media sites. Reference [19] studies the sentiments of McDonalds and KFC to appear which food chain has more fame.

Custom-based models use Twitter analysis case studies based on topic, function, geographic area, and/or any suitable Twitter features as applicable. The Twitter analysis may have an impact on human life, including students’ and patients’ life. Twitter is used to study diabetes management [2]. Thom and Kruger investigated the possibility of Twitter in life saving scenarios during the 2013 German Floods. The protest hashtag #BlackLivesMatter has emerged after the killing of an American Black person by a White one. The counter-protest hashtag #AllLivesMatter has emerged to show equal attention to lives regardless of race [15]. The study of [15] focused on analyzing the divergence of Twitter topics using two major hashtags (i.e., #BlackLivesMatter and #AllLivesMatter). The study [16] uses the actor-network theory, which emphasizes the importance of a user based on several connections. Thereby, they provide a co-occurrence matrix of discussed topics (as hashtags) and related Twitter users. Some researchers have studied initiatives taken by companies on Twitter features as applicable. The Twitter analysis may have reliability of public opinion [10]. Therefore, it is not only analyzing the current topics but moreover, anticipate the patterns and results of few acute real-world events such as gun violence and drug misuse. The emotional topics are not only capable of providing insights in community discussions, but also comprise of the machine learning algorithms that analyze Twitter content form linguistic and psychological perspectives [13]. Most machine learning models aim to identify and verify Twitter users [14]. This research analyzes discussions about a topic to gain insights into it. Twitter has been utilized for investigating opinions (sentiments) on numerous social media sites. Reference [19] studies the sentiments of McDonalds and KFC to appear which food chain has more fame.

The metric-based approaches adopt methods from network science. They use centralization concepts as well as modularity algorithms to extract results. The centralization is a measure node’s importance (a node is a user or a tweet) concerning the centrality of other nodes [18], [19]. There are several metrics for centralization that are used in this research, the common ones include degree, betweenness, and PageRank. The degree measures the number of links coming to (in-degree) or departing from a specific node (out-degree). These measures can be used to measure the influence of a user. It reveals who were the most active users.
The betweenness-centrality measures the effectiveness of a Twitter user, usually based on the minimum number of edges (called the shortest path) [18]. Therefore, in any network, the innovators and brokers are considered to be the people with high Betweenness [20]. PageRank ranks users based on their importance in the network. In addition to centrality measures, modularity based algorithms identify communities (or clusters) in a network [21].

Network science has many applications like in: community detection [22], [23], influence propagation [24], [25], event detection [26]–[29], user characterization [30], [31], education [3],[32], linguistics [33], sentiment analysis [34], [35], analyzing money laundering [36], software engineering [37], altmetrics [38]–[40], archeology [41], disaster management [42], tourism [43], socio-economics [4], and advertising [44].

III. METHODOLOGY

In this research we use methods from network science and collectively group them as a method called “Understanding Public Discussions Using Network Science” (UPDUNS). The methods used in UPDUNS already exist in the literature of network science, however, according to the best of our knowledge, these methods have not yet been used for studying the public discussions about “Saudi Women Can Drive”. We use UPDUNS to understand the public discussions about #SaudiWomenCanDrive as depicted in Figure 1. The methodology consists of the following main components: data collection, network extraction, preprocessing, network analysis, visualization, and evaluation, detailed in the subsequent section.

A. DATA COLLECTION

The first step is collecting tweets. The data is gathered using Twitter’s free Application Programming Interface (API). There are different methods provided by the Twitter API to collect data. Two methods: search and filter are mostly used by researchers for collecting data. Search API is used to get data for the past 5 to 10 days, whereas filter API allows to collect data in real-time. As we analyze the public discussions centered around the topic of allowing women to drive in Saudi Arabia and the decision in support of allowing women to drive was made on September 26, 2017 by King Salman of the Kingdom of Saudi Arabia. Twitter’s Search API allowed us to collect data about the discussion four days post announcement. We collected all the tweets having one of the hashtags mentioned in Table 1. We only used the hashtags written in Arabic to focus on the tweets posted in Arabic. We used terms which were in favor of the movement, against it, and neutral. The tweets are gathered from the 26th of September 2017 till the 30th of September 2017. Consequently, the dataset contains around two million tweets, and 630,000 users. The networks extracted are available on GitHub.3

3https://github.com/rabeehabbasi/SaudiWomenCanDrive

B. NETWORK EXTRACTION

For large networks, non-effective users should be ignored during the analysis. One preliminary possibility for analysis is to use the available information in its core state to reveal the most active users. These most-active users are estimated by the high centrality measures like degree centrality, PageRank, and betweenness centrality [46]. We use the value of \( p = 0.85 \) for PageRank, as widely used in the literature. Within the course of the investigation, the initial information set is converted into four networks: Mention-Network, Co-Mention-Network, Hashtag-Network, and Retweet-Network, interested readers can find the minuiae in our earlier work [47]. The substance of these networks complies involving network science for succeeding analysis and formulation. Table 2 shows the network science measures used in this study. Table 3 lists the statistics of each network extracted from public discussions. Nodes and edges in each of these networks are formed as follows:

- **Retweet-Network**: Nodes represent users, an edge is formed from user 1 towards user 2 (directed), if user 1 retweets user 2.
- **Mention-Network**: Nodes represent users, an edge between user 1 and user 2 (directed) represents that user 1 has mentioned user 2 in at least one of his tweets.
- **Co-Mention-Network**: Nodes represent users. There is an edge between two users (undirected), if both of them are mentioned in the same tweet. If multiple users are mentioned in a tweet, they form a clique.
- **Hashtag-Network**: Nodes represent hashtags and there is an edge between two hashtags (undirected), if they appear together in the same tweet. If multiple hashtags appear in a tweet, they form a clique.

C. NETWORK ANALYSIS AND VISUALIZATION

The network analysis of the proposed model is based on the extracting network knowledge based on centrality,
modularity to detect communities and Page Rank to detect influencers. All these metrics are conducted all over the four types of networks, discussed above. The appropriate metrics are considered to find out the results for the investigated research. Finally, data is visualized in a network analysis tool Gephi [48].

### TABLE 1. Hashtags used for Data Extraction.

| Hashtag                | Translation       |
|------------------------|-------------------|
| #الملك ينصر قيادة المرأة | The king supports women’s driving |
| #الشعب يرفض قيادة المرأة | Saudis reject women’s driving |
| # السماح للمرأة بالقيادة | Allowing women to drive |
| #السعودية              | Saudi Arabia      |

### TABLE 2. Network Science Measures used in UPDUNS.

| Measure     | Network  | Purpose                                                                 |
|-------------|----------|-------------------------------------------------------------------------|
| Betweenness | Mention  | The centralization is a degree of the significance of a node (a user or a tweet) in relation to how the centrality of other nodes [18]. The innovators and brokers in any network are reflected to be the people with high Betweenness [20]. |
|             | Co-Mention Hashtag Retweet |                                                                              |
| Degree      | Mention  | The centralization of a Twitter user can be indicated by in-degree or out-degree. It reveals who are the most active users. |
|             | Co-Mention Hashtag Retweet |                                                                              |
| PageRank    | Hashtag  | Google’s PageRank rank websites based on its importance. Tweets that use hashtags are an example of the PageRank metric. |
| (p=0.85)    |          |                                                                         |
| Modularity  | Mention  | Modularity measures how network information is clustered in a network (modules, communities, or groups) [45]. Therefore, for Twitter network, it captures the extent to which groups of people are disconnected from each other; therefore, people have their community and topics of interest. A high modularity score indicates an erudite interior structure, which is commonly referred to as the community structure. In the current research context, there are two major communities: opponents and supporters about the decision of a woman driving a car. |
| (Louvain, resolution=1) | Retweet |                                                                              |

### FIGURE 2. Communities in the co-mention network.

**D. EVALUATION**

The huge volume of data and the absence of any benchmark data make the automated evaluation challenging. Therefore, expert judgments have been used for evaluating the results. To acquire ranking, analysis is conducted conforming the correctness of top-k results. Furthermore, the qualitative analysis has been ensured to conform the meaningfulness of the communities.

### E. SOCIAL MEDIA AND NETWORK ANALYSIS TOOLS

There is a variety of tools available for social media and network analysis like NodeXL and Gephi, and many programming libraries like NetworkX and iGraph. These tools and libraries can be used to analyze a network like we did in this research, however, none of these tools automatically provide the qualitative and in-depth analysis presented in this research.

### IV. RESULTS AND DISCUSSIONS

**A. COHESIVE GROUPS IDENTIFICATION**

The identification of the cohesive groups of people in the discussion is an extension to our previous work [47], where identification of the critical people in a discussion is studied. In network science, a community can be termed as a collection of nodes that have many edges among themselves. Therefore, groups presume the fact that the information within the group flows much faster compared to information with the outer nodes. Finding communities, commonly referred to as modularity, is often visualized and processed by highly efficient algorithms [49].

One of the most common practices when dealing with modularity is to analyze it together, along with betweenness centrality. The justification to this approach is the assumption of nodes that connect diverse groups are crucial “switches”
that smooth information flow, and consequently reveals the “vulnerability” of a network structure. Therefore, we use it for community detection. One can identify the effect of removing a bridge participant between two groups, including network change and alternative connectors. The communities in the current work are detected using undirected Co-Mention-Network.

We used Louvain algorithm with default resolution value of 1 to extract communities. Louvain is a scalable community detection algorithm and can extract communities in very large networks very fast [49]. The Co-Mention-Network of the studied dataset has 1298 communities, and modularity of 0.81. Figure 2 visualizes the biggest five communities in this network. These communities are individually discussed in the following paragraphs. In each of these individual communities, top users are ranked based on their betweenness centrality.

The first community represents 5.11% nodes of the whole network and contains 300 nodes and 661 edges. Figure 3 and associated Table 4 represent the top users with respect to betweenness in the largest community in the Co-Mention-Network. Figure 3 shows a strong relationship between LoujainHathloul and manal_alsharif as strong women driving activists. Most of the users in this community are women driving activists and social media figures. Furthermore, most of the tweets that mention these accounts provide thankfulness and offer congratulations regarding the positive outcome of the decision. This community contains the supporter.

The third community represents 3.8% nodes of the whole network, contains 223 nodes and 425 edges. Figure 5 and associated Table 6 represent the top users in the third community with respect to betweenness in the Co-Mention-Network. This community contains some government accounts and most of the tweets that co-mention these accounts discuss the practical aspects and regulations around women receiving the right to drive.

The fourth community represents 2.31% nodes of the whole network, contains 136 nodes and 225 edges. Figure 6 and associated Table 7 represent the top users with respect to betweenness in the fourth community in the co-mention network.

Uber and Careem are the most popular companies that offer rides in Saudi Arabia. Their account has a thick edge between them as they co-mention together. CNNArabic user has an edge with Careem since it mentions Careem to praise the job opportunities

| Table 4. Top Users in the First Community With Respect to Betweenness |
|-----------------------------|-----------------------------|
| User                        | Description                                                                 |
| LoujainHathloul              | Among the primary promoters of the decision                                   |
| AhlamAlShamsi                | Ahlam is an Emirati singer, who congratulated Saudi women on the Women Driving decision |
| LatifahAshaalan              | Writer and associate professor of psychology at Princess Nora University, also a member of the Consultative Assembly of Saudi Arabia. She is a well-known supporter of this cause |
| manal_alsharif              | Among the primary promoters of the decision                                    |
| Naifco                       | Naif Madkhali, social media figure                                             |
| b_b12331                     | Active user, shared many tweets to support the decision                        |
| fdeet_alnssr                 | Naif bin Osaid, social media figure                                            |
| MeshariGhamdi                | Active user, shared many tweets to support the decision                        |
| DAHmadq84                    | Active user, shared many tweets to support the decision                        |
| MERiAM_AL3TEEIBE             | Active user, shared many tweets to support the decision                        |

| Table 5. Top Users in the Second Community With Respect to Betweenness |
|-----------------------------|-----------------------------|
| User                        | Description                                                                 |
| KingSalman muhammad_1992   | “King Salman of KSA”                                                   |
| M_Naif_Alsaoud              | “Active user, opponent”                                                 |
| ArabianVeritas              | “Supervision of products conformity: Verification of Conformity PROGRAMME. Conformity to Saudi Standards (SASO) and Technical Regulations” |
| TahrirLive                  | “Active twitter account”                                                 |
| Alkingjuber                 | “Active user, opponent”                                                  |
| eew38635201                 | “Active user”                                                            |
| SaudiVision2030             | “The official account for the vision program of Saudi Arabia”             |

The fourth community represents 2.31% nodes of the whole network, contains 136 nodes and 225 edges. Figure 6 and associated Table 7 represent the top users with respect to betweenness in the fourth community in the co-mention network.

Uber and Careem are the most popular companies that offer rides in Saudi Arabia. Their account has a thick edge between them as they co-mention together. CNNArabic user has an edge with Careem since it mentions Careem to praise the job opportunities
The second community represents 4.34% nodes of the whole network, contains 255 nodes and 292 edges. Figure 4 and associative Table 5 represent the top users in the second community in the co-mention network with respect to betweenness. This community contains the decision-maker, King Salman. There is only one edge between King Salman and Prince Muhammad bin Nayef. It is a tweet that co-mentions both and is related to the decision’s declaration. User Abdullah_50560, Alkingjaber and muhammad__1992 are considered as opponents and these users have an edge with King Salman, due to retweeting many tweets that disagreed with the decision. There is an edge between SaudiVision2030 and KingSalman which highlights that the women are a vital part of Saudi development.

This community contains prominent Islamic figures and Council of Senior Scholars that supported the King’s decision.

B. BRIDGE IDENTIFICATION

In any network, it is important to identify the nodes which play the roles of bridges, i.e., connecting various communities in the network. The bridging role is calculated using the betweenness centrality. Table 9 shows that the top users with the highest betweenness centrality in the co-mention network. Furthermore, it designates the top users who get co-mentioned due to their influential power. Therefore, such users are considered as the major players in the discussion.

Figure 8 shows the top users, according to betweenness centrality, where the dark color and size refers the higher betweenness node. King Salman has a co-mention with most of the users in the network. KingSalman has an edge with MOISaudiArabia, since the decision is related to the interior affairs, the MOISaudiArabia is co-mentioned with KingSalman to have the attention towards this issue. The reason is that The Ministry of Interior is responsible for organizing interior affairs and servicing citizens. Explanation of one of the tweets that mention them is as following:

According to @KingSalman’s order, @MOISaudiArabia to draft necessary amendments to the traffic regulations #SaudiWomenCanDrive

provided by Careem for women. AboutHerOFCL: share some tweets about the opportunities provided by Uber, Careem. StateDept: welcomed the decision.

The fifth community represents 2.13% nodes of the whole network, contains 125 nodes and 165 edges. Figure 7 and associated Table 8 represent the top users with respect to betweenness in the 5th community in the co-mention network.

FIGURE 4. Connections among the top users of the second community.

TABLE 6. Top Users in the Third Community with Respect to Betweenness.

| User          | Description                                                                 |
|---------------|-----------------------------------------------------------------------------|
| MOISaudiArabia| “The Ministry of the Interior of Saudi Arabia”                               |
| saudq1978     | “Saud bin Abdullah al-Qahtani is a Saudi Arabian consultant and ex-advisor to royal court. He served as a legal advisor to King Abdullah of Saudi Arabia and a media consultant to Crown Prince Mohammed bin Salman before his dismissal in late 2018.” |
| eMoroor       | “The official account of the Directorate of Traffic”                          |
| YSulfaisal    | The official account for the Ministry of the Interior’s app KollonaAmm, which “empowers all citizens and inhabitants in Saudi Arabia to mimic the function of a police officer” by recording and submitting reports |
| Kammapp       | The official account for the Ministry of the Interior’s app KollonaAmm, which “empowers all citizens and inhabitants in Saudi Arabia to mimic the function of a police officer” by recording and submitting reports |
| DAhmadq84     | “Active user, supporter and shared many tweets to support the decision”       |
| Allahim       | “Active user, supporter and shared many tweets to support the decision”       |
| Badermasaker  | “The head of the Private Office of HRH. The Crown Prince and Chairperson of the Misk initiative center” |
| TurkiAldakhil | “The Saudi ambassador to the United Arab Emirates”                            |
| bip_ksa       | “The accredited account of the Kingdom of Saudi Arabia’s Public Prosecution,” |

FIGURE 5. Connections among the top users of the third community.
According to the Figure 8, there are edges between Alwaleed_Talal, CNNArabic, and KingSalman. The Alwaleed_Talal account is an account of the famous Saudi prince businessperson, and one of the influential people in Saudi Arabia. He published a report in 2016 to stop the debate and asking for allowing Saudi women to drive. CNNArabic is a news website and part of the CNN network. The CNNArabic mentioning Alwaleed_Talal and published its report about women driving. The ssa_at engages in the issue because they agreed and supported the decision. Notably, there is a community of women supporters at the bottom of the network.

Table 10 describes the mention network sorted by betweenness centrality. In the derived filtered network, Abdullah_50560 is mentioned more than that he mentioned other entities, as indicated by his 228 in-links and 57 out-links. The investigation showed that he had around 77 followers at the beginning of this conversation. However, later, primarily due to his activity on this topic, his followers increased to 144 (this is from Sep 27 to Oct 6th, 2017).
TABLE 9. Top Ten Users in the Co-Mention Network, Ranked based on Betweenness Centrality.

| User               | Degree | Betweenness (x1000) |
|--------------------|--------|---------------------|
| KingSalman         | 286    | 1,343.7             |
| LoujainHathfoul    | 111    | 397.1               |
| KSA24              | 66     | 285.7               |
| Algassabinasser    | 87     | 271.2               |
| ssa_at             | 66     | 235.6               |
| fayez_malki        | 57     | 233.8               |
| SaudNews50         | 70     | 224.2               |
| MOISaudiArabia     | 68     | 218.8               |
| CareemKSA          | 51     | 207.6               |
| saudq1978          | 43     | 184.9               |

TABLE 10. Top Ten Users in the Mention Network, Ranked based on Betweenness Centrality.

| User            | In-Degree | Out-Degree | Betweenness (x1000) |
|-----------------|-----------|------------|---------------------|
| Abdullah_50560  | 228       | 57         | 598.5               |
| alkinha505      | 5         | 119        | 331.7               |
| mfabe_50        | 59        | 40         | 157.3               |
| K700G7          | 15        | 11         | 137.2               |
| Y_Alamri96      | 13        | 22         | 129.6               |
| soooso99991     | 22        | 6          | 127.5               |
| nafi20071       | 4         | 24         | 119.3               |
| asd025056       | 3         | 78         | 118.0               |
| pmYukfonkRQrxM  | 9         | 20         | 116.9               |
| Sn57562993      | 2         | 24         | 110.4               |

TABLE 11. Top Ten Users in the Retweet Network, Ranked based on Betweenness Centrality.

| User            | In-Degree | Out-Degree | Betweenness (x1000) |
|-----------------|-----------|------------|---------------------|
| Abdullah_50560  | 375       | 55         | 1,176.7             |
| alkinha505      | 5         | 101        | 658.6               |
| K700G7          | 20        | 11         | 572.8               |
| Y_Alamri96      | 18        | 21         | 542.6               |
| loody_ahlial     | 37        | 2          | 491.4               |
| sooso99991      | 41        | 4          | 486.9               |
| rhofa           | 227       | 18         | 458.0               |
| Rowan2045       | 35        | 4          | 454.0               |
| Klatish         | 161       | 4          | 449.3               |
| nafi20071       | 5         | 24         | 448.3               |

TABLE 12. Statistics of the Hashtag Network.

| Statistic        | Measure |
|------------------|---------|
| Number of nodes  | 14,305  |
| Number of edges  | 59,638  |
| Average degree   | 8.34    |
| Average Weighted Degree | 194.29 |
| Network Diameter | 9       |
| Density          | 0.001   |
| Modularity       | 0.64    |
| Communities      | 152     |

he appears as the most central in terms of betweenness. When the investigations of a sample of his activity is conducted, it revealed that the user is sharing tweets against the decision.

Another user of interest is alkinha505, who is having only 44 followers, has mentioned more users while he is mentioned only by a few. Therefore, the score of in-degree versus out-degree analysis results to “discard” this user considering not very influential. However, in terms of bridging activity, his role is vital as he is the second-highest user according to betweenness centrality. Therefore, it was deduced that the most productive players in central groups are Abdullah_50560 and alkinha505 being considered as an opponent to the decision. Rest of the nodes are marginal with respect to in/out degrees in relation with betweenness centrality. It suggests that this so-called “marginal” activity, against all the odds, connects distant regions in the network.

Table 11 shows the ten highest scores of betweenness in the Retweet-Network. The rank of users of the first two users is like those shown in Table 10; however, the user K700G7 is at third place. The justification is that the nodes are rarely retweeted. The nodes are also retweeted themselves infrequently (at least, compared to the massive players in the network). Nonetheless, K700G7 appears as a necessary connector within diverse groups.

C. TOPIC ANALYSIS

It is critical to analyze the extent to which people debate the common interests (or topics) to understand their interests and future needs or complaints. In identifying topics of debate, the hashtag network is a revealing tool, due to the reason that hashtags commonly appear in the same tweet. The fact that the hashtags are commonly accepted as short descriptions of the stance of the Twitter space, consequent analyzing them which provides different facets of the underlying conversation.

Table 12 shows the statistics for the network, and associated Figure 9 shows the hashtag network. The colors represent topical communities. The network’s diameter is lower than the network diameter of the Retweet-Network and Mention-Network, which signifies the connectivity of the internal structure of the network. The average degree of 8.3 indicates that the users are more tightly connected as compared to the Retweet-Network and Mention-network. The top ten nodes (hashtags) with the highest degrees are shown in Table 13 and relations between them are shown in Figure 10.

It is observed from the hashtag names that many of them are similar. For example, the hashtags 1, 3, 7, and 10 are the same, but since the Arabic language has a variety of ways to spell the same word due to presence or absence of vowels (tashkeel) and use of single or multiple underscores (_) for separating words.

It is further observed that there is a strong association between the hashtags “The King supports women’s driving” and “Allowing women to drive”, which shows that these two hashtags have appeared in many tweets together.

Another strong association is between the hashtags “The King supports women’s driving” and “Saudi Arabia”. These hashtags often appear together, as the foremost relates to the decision, and then later mention Saudi Arabia. Identifying
the most common hashtags could be acquired comfortably with the tools of core network analysis too. Therefore, these hashtags are considered as most influential topics related to the public discussions.

Figure 11 demonstrates the heatmap of the hashtag co-occurrence, with multiple similar hashtags merged. Each darker spot indicates the two hashtags appeared together, usually with much weight, since these are the most encountered hashtag pairs in the data.

Distinct types of networks provide different perspectives into the discussions, for example, mention and retweet networks help in finding influential people in the discussions, as many other people endorse their tweets through mentions and retweets. Similarly, co-mention network helps in identifying cohesive group of people, as they are often mentioned together in the tweets. Mention, retweet, and co-mention networks are associated with people related to the public discussion, however, if one wants to explore the topics being discussed, a hashtag network provides such insights, not only hashtag network provides most important topics, but also the associations among these topics.

### TABLE 13. Top Ten Hashtags with Respect to Degree.

| Hashtag                                                                 | Degree | Betweenness (<1000) |
|------------------------------------------------------------------------|--------|---------------------|
| الملك ينصّر للقيادة المراه  | 5,046  | 39,246.0            |
| (The King supports women’s driving)                                   |        |                     |
| الشعب يرفض نقص القيادة المراه  | 2,926  | 19,610.1            |
| (Saudis refuse women’s driving)                                       |        |                     |
| الملك ينصّر للقيادة المراه  | 2,686  | 17,516.5            |
| (The King supports women’s driving)                                   |        |                     |
| SaudiWomenCanDrive                                                    | 1,587  | 12,206.5            |
| الشعب يرفض قيادة المرأة 2  | 1,405  | 8,135.3             |
| (Saudis refuse women’s driving2)                                      |        |                     |
| السعودية                                                              | 1,228  | 3,946.7             |
| (Saudi Arabia)                                                        |        |                     |
| الملك ينصّر قيادة المرأة  | 949    | 3,980.4             |
| (The King supports women’s driving)                                   |        |                     |
| السماح للمرأة بالقيادة  | 877    | 3,314.2             |
| (Allowing women to drive)                                             |        |                     |
| مبروك للنساء الوطن                                                     | 853    | 2,063.7             |
| (Congratulations Saudi Women)                                         |        |                     |
| الملك ينصّر قيادة المرأة  | 746    | 2,720.6             |
| (The King supports women’s driving)                                   |        |                     |

**Figure 9. The hashtag network.**

**Figure 10. Connections among the top 10 hashtags.**

**Figure 11. Heatmap of hashtags showing which hashtags appear with others most frequently.**
V. CONCLUSION
This paper analyzes public discussions about allowing women to drive in the Kingdom of Saudi Arabia using methods from network science. Four distinct types of networks: retweet-network, mention-network, hashtag-network, and co-mention-network are used for the analysis. Each network and associated suitable measures help in identifying influential people, communities involved in these discussions, and the topics discussed. The relationships among the most influential people and popular topics are also analyzed. The acquired results allow us to understand the diverse communities (like activists, political leadership, government institutions, news agencies) formed around the public discussion. We identify influential individuals from varying backgrounds and how they relate with other. Using the methods from network science, we can identify people like King Salman who play influential role of bridging diverse communities.

This research has used methods from network science to get results which are further qualitatively analyzed. Most of the case studies which use network science methods, follow a similar approach. This is in contrast with the research being conducted in machine learning, where results are quantitatively measured in form of accuracy measures and can be compared with other methods. Like other research works which use similar methodology, the results presented in this research are difficult to evaluate in absence of a ground truth (which is available in supervised machine learning based research). However, by following the methodology used in this research on a dataset about a public discussion, one should expect to get similar results, but not be exactly the same. The networks used in this research can be downloaded from GitHub. This research has analyzed a single discussion in Arabic, in future we plan to analyze a broader set of public discussions happening around the world in various languages.

VI. DECLARATION
This research has not been funded by any agency and there is no conflict of interest.

REFERENCES
[1] R. A. Abbasi, O. Maqbool, M. Mushtaq, N. R. Aljohani, A. Daud, J. S. Alowibdi, and B. Shahzad, “Saving lives using social media: Analysis of the role of Twitter for personal blood donation requests and dissemination,” Telematics Inform., vol. 35, no. 4, pp. 892–912, Jul. 2018.
[2] M. L. Litchman, D. Lewis, L. A. Kelly, and P. M. Gee, “Twitter analysis of #OpenAPS DIY artificial pancreas technology use suggests improved A1C and quality of life,” J. Diabetes Sci. Technol., vol. 13, no. 2, pp. 164–170, 2019.
[3] R. Isba, K. Woolf, and R. Hanneman, “Social network analysis in medical education,” Med. Educ., vol. 51, no. 1, pp. 81–88, Jan. 2017.
[4] J. Gao, Y.-C. Zhang, and T. Zhou, “Computational socioeconomics,” Phys. Rep., vol. 817, pp. 1–104, Jul. 2019.
[5] J. P. Alperin, C. J. Gomez, and S. Haustein, “Identifying diffusion patterns of research articles on Twitter: A case study of online engagement with open access articles,” Public Understand. Sci., vol. 28, no. 1, pp. 2–18, Jan. 2019.
[6] A. Alotaibi, “Why the panic? Gendered moral panics and the Saudi ban on women driving,” George Washington Univ., Washington, DC, USA, 2017.
[7] E. Alhussein, “Triangle of change: The situation of women in Saudi Arabia,” Norwegian Peacebuilding Resource Centre (NOREF), Norway, Expert Anal. Rep., 2015.
[8] K. Krane and F. Majid, “Women driving in Saudi Arabia: Ban lifted, what are the economic and health effects?” Rice Univ. Baker Inst. Public Policy, Houston, TX, USA, Issue Brief no. 06.13.18, 2018.
[9] I. Himelboim, M. A. Smith, L. Rainie, B. Shneiderman, and C. Espina, “Classifying Twitter topic-networks using social network analysis,” Soc. Media Soc., vol. 3, no. 1, 2017, Art. no. 205630517691545.
[10] U. Kursuncu, M. Gaur, U. Lokala, K. Thirunarayan, A. Sheth, and I. B. Arpinar, “Predictive analysis on Twitter: Techniques and applications,” in Emerging Research Challenges and Opportunities in Computational Social Network Analysis and Mining, Cham, Switzerland: Springer, 2019, pp. 67–104.
[11] A. Addawood, A. Alshamrani, A. Alqahtani, J. Diesner, and D. Broniatowski, “Women’s driving in Saudi Arabia—Analyzing the discussion of a controversial topic on Twitter,” in Proc. Int. Conf. Social Comput., Behav.-Cultural Modeling, Predict. Behav. Represent. Modeling Simulation (BRiMS), 2018.
[12] L. Altoaimy, “Driving change on Twitter: A corpus-assisted discourse analysis of the Twitter debates on the Saudi ban on women driving,” Social Sci., vol. 7, no. 5, p. 81, May 2018.
[13] C. J. Hutto and E. Gilbert, “VADER: A parsimonious rule-based model for sentiment analysis of social media text,” in Proc. 8th Int. Conf. Weblogs Social Media (ICWSM), 2014, pp. 216–225.
[14] O. Halvani, L. Grauer, and I. Vogel, “Authorship verification in the absence of explicit features and thresholds,” in Proc. Eur. Conf. Inf. Retr., 2018, pp. 454–465.
[15] R. J. Gallagher, A. J. Reagan, C. M. Danforth, and P. S. Dodds, “Divergent discourse between protests and counter-protests: #BlackLivesMatter and #AllLivesMatter,” PLoS ONE, vol. 13, no. 4, 2018, Art. no. e0195644.
[16] I. Hellsten and L. Leydesdorff, “Automated analysis of actor–topic networks on Twitter: New approaches to the analysis of socio-semantic works,” J. Assoc. Inf. Sci. Technol., vol. 71, no. 1, pp. 3–15, Jan. 2020.
[17] B. Alotaibi, R. A. Abbasi, M. A. Aslam, K. Saeed, and D. Alahmadi, “Startup initiative response analysis (SIRA) framework for analyzing startup initiatives on Twitter,” IEEE Access, vol. 8, pp. 10718–10730, 2020.
[18] L. C. Freeman, “A set of measures of centrality based on betweenness,” Sociometry, vol. 40, no. 1, pp. 35–41, Mar. 1977.
[19] L. Liu, D. Chen, X.-L. Ren, Q.-M. Zhang, Y.-C. Zhang, and T. Zhou, “Vital nodes identification in complex networks,” Phys. Rev., vol. 650, pp. 1–63, Sep. 2016.
[20] U. Brandes, “A faster algorithm for betweenness centrality,” J. Math. Sociol., vol. 25, no. 2, pp. 163–177, 2001.
[21] M. Komorowski, T. D. Hsu, and N. Deligiannis, “Twitter data analysis for studying communities of practice in the media industry,” Telematics Inform., vol. 35, no. 1, pp. 195–212, Apr. 2018.
[22] A. Said, R. A. Abbasi, O. Maqbool, A. Daud, and N. R. Aljohani, “CGA: A clustering coefficient based genetic algorithm for detecting communities in social networks,” Appl. Soft Comput. J., vol. 63, pp. 59–90, Feb. 2018.
[23] W. Silva, Á. Santana, F. Lobato, and M. Pinheiro, “A methodology for community detection in Twitter,” in Proc. Int. Conf. Web Intell., Aug. 2017, pp. 1006–1009.
[24] N. E. Friedkin, A. V. Proskurnikov, R. Tempo, and S. E. Parsyev, “Network science on belief system dynamics under logic constraints,” Science, vol. 354, no. 6310, pp. 321–326, Oct. 2016.
[25] N. Alnujui, A. Datta, and J. Li, “Influence propagation model for clique-based community detection in social networks,” J. Grid Comput., vol. 13, no. 4, 2014, Art. no. e0195644.
[26] Z. Saeed, R. A. Abbasi, I. Razzak, O. Maqbool, A. Sadaf, and G. Xu, “Enhanced heartbeat graph for emerging event detection on Twitter using time series networks,” Expert Syst. Appl., vol. 136, pp. 115–132, Dec. 2019.
[27] Z. Saeed, R. A. Abbasi, O. Maqbool, A. Sadaf, I. Razzak, A. Daud, N. R. Aljohani, and G. Xu, “What’s happening around the world? A survey and framework on event detection techniques on Twitter,” J. Grid Comput., vol. 17, no. 2, pp. 279–312, Jun. 2019.
[28] Z. Saeed, R. A. Abbasi, A. Sadaf, M. I. Razzak, and G. Xu, “Text stream to temporal network—A dynamic heartbeat graph to detect emerging events on Twitter,” in Proc. Pacific–Asia Conf. Knowl. Discovery Data Mining, in Lecture Notes in Computer Science, vol. 10938, 2018, pp. 534–545.

4https://github.com/rabeechabbi/SaudiWomenCanDrive
ZUBAIDA JASTANIA received the master’s degree from the Department of Information System, King Abdulaziz University, Saudi Arabia, in 2020. Her research interests include data analytics and social network analysis.

MUHAMMAD AHTISHAM ASLAM received the master’s degree in computer science from Hamdard University, Lahore, Pakistan, in 2002, and the Ph.D. degree from the University of Leipzig, Leipzig, Germany, in 2007. From 2003 to 2007, he was worked as a Research Assistant with the University of Leipzig. He is currently working as an Associate Professor with the Department of Information Systems, King Abdulaziz University, Jeddah, Saudi Arabia. He has been working as a Senior Staff Researcher with the Artificial Intelligence Laboratory, Knowledge Technology Cluster, Malaysian Institute of Microelectronic Systems (MIMOS), Kuala Lumpur, Malaysia. He also has been working as an Assistant Professor with the COMSATS Institute of Information Technology, Lahore. His research interests include semantic web, linked open data, data science, and social networks.

TARIQ JAMIL SAIFULLAH KHANZADA received the Ph.D. degree in the field of wireless information and communication engineering from the OvGU University of Magdeburg, Germany, in 2010. Since September 2019, he has been working with the IS Department, Faculty of Computing and IT, KAU Jeddah, KSA, as a Professor. He is on lien from the Department of CS Engineering, Mehran UET Jamshoro Pakistan, where he has been working as a Professor (bps-21), since October 2016, and worked for various academic positions, since 2001. He got various national and international meritorious awards and research academic fundings. He had established the Muhammad Institute of Science and Technology MirpurKhas Pakistan, as a Pioneer Head of faculty. He has more than 50 international conferences and journal publications, including a coauthorship of a book on computers. His research interests include wireless indoor positioning systems, OFFDM & SCT transceiver simulations, signal processing for wireless communications, data mining/data sciences methods, and applications of AI algorithms.

KHAWAJA MOYEZULLAH GHORI received the master’s degree in computer sciences from the National University of Computer and Emerging Sciences, Pakistan, and the Ph.D. degree in computer science from the University of Debrecen, Hungary. He has been working as an Assistant Professor with the Department of Computer Science, National University of Modern Languages, Islamabad, since 2004. His research interests include machine learning, data mining, deep learning, and neural networks.

Z. Jastania et al.: Analyzing Public Discussions About #SaudiWomenCanDrive Using Network Science