Impressions of the Community of Makkah on the Hajj in the Light of Covid-19 Pandemic: Quantitative and AI-based Sentiment Analyses

Hani A. Aldhubaib*

Department of Electrical Engineering, College of Engineering and Islamic Architecture, Umm Al-Qura University, Makkah, Saudi Arabia

Abstract. Throughout history, the community of Makkah has been known to make every effort to serve the Muslims who come to perform the Hajj ritual (Pilgrimage), the fifth pillar of Islam. However, the Hajj of 1441 H (2020) was different due to the outbreak of the new Coronavirus (Covid-19). This effect caused numerous changes in many aspects of human life in various societies, including the community of Makkah. While the people of Makkah were preparing to provide what they used to do every year to support and facilitate services for pilgrims, they had to follow the instructions of the Ministry of Health to implement many precautionary and preventive measures to avoid infection of Covid-19; most important of which is staying home. This study mainly aims to observe and analyze social-related changes that may be present or absent for the first time in the community of Makkah during the Hajj of 1441 H. Two main methods were used: AI-based sentiment analysis for a sample of tweets as well as a questionnaire for a random sample of Makkah residents.

Keywords: Covid-19, Hajj, sentiment analysis, Twitter

1. Introduction

Since late 2019, the outbreak of Covid-19 has had adverse effects on almost all sectors, including economic, health, educational and social [1], which significantly impacted many aspects of life, including the Hajj. In the Hajj of 1441 H, a limited number of pilgrims were allowed to perform the Hajj rituals, adhering to the instructions issued by the Ministry of Health following specific preventive and precautionary protocols to prevent the spread of the Covid-19. This incident had a significant effect on the feelings of individuals who found social media platforms, particularly Twitter, as free spaces to express their emotions, impressions, and opinions. Several techniques have been used in the literature to investigate the impression and feelings of individuals in social media, one important of which is sentiment analysis.

Recently, there has been a growing interest in sentiment tracking and analysis of news and commentary on various social media platforms, which has targeted several subjects of human life, including, for instance, health [2]–[4], disasters [5], [6], tourism [7], [8], political [9], [10]. Analysis of this data helps understand the behavior and thoughts of individuals and society about specific events and issues and predict future trends. Many studies in the literature have discussed numerous topics related to tracking and sentiment analytics; nevertheless, research studies addressing this subject in Arabic have been relatively modest compared to the English language [11]–[13].

This paper studies the feelings and impressions of the Makkah residents towards the Hajj of 1441 H (2020), considering the changes associated with the Covid-19 pandemic. The data was collected using two primary methods: Twitter and questionnaires. More than 14 thousand Arabic tweets were collected and classified according to positive, neutral, and negative classes using the sentiment analysis (SA) method described in section 3. Finally, the impressions of the Makkah community were obtained and discussed according to several issues closely related to the Hajj and Eid days. These Islamic festival days begin on the tenth day of Dhu Al-
Hijjah month (the twelfth and last month of the Islamic Calendar).

The remainder of the paper is as follows: Section 2 provides a brief review of the related studies. Section 3 describes the methodology used in this paper, and a discussion of the results is given in section 4. Finally, Section 5 concludes this paper.

2. Related literature

Social media has become very popular and fast-growing globally, such as Twitter, Facebook, and Instagram. Due to the technological-related evolution that changed how information is exchanged globally, the public uses social media platforms to express their opinions. Although many challenges are associated with social media data, social media has become the primary source of obtaining and the most influential in forming public opinion. Several reports and studies discussed this issue in the literature, including [14]–[17]. In the fourth quarter of 2021, the average monetizable daily active users (mDAU) reached 217 million worldwide [18]. In Saudi Arabia, about 9.9 million people use Twitter regularly, ranking the fourth-highest usage percentage in the world [19]. Twitter has become one of the most popular social networking sites that allow users to express their feelings, opinions, and ideas. Many other uses have enabled users to benefit from Twitter in different aspects, for example, education [20], managing business [21], socialization [19], and others [22]–[24]. Given the great importance of Twitter in societies, many researchers and scientists have used the Twitter platform to conduct many studies and research in various sciences. Many research attempts dealt with Hajj and Umrah system applications from other disciplines and research areas [25]–[28]. The literature survey of this paper addresses studies that focused on the use of social media in general and Twitter in particular to support and develop the Hajj and Umrah system.

Al-Sadiq and Ahmed [29] studied the attitudes and motives of pilgrims for using the social media websites of the General Presidency of Haramain (GPH). The authors used a questionnaire distributed among a sample of 400 pilgrims who use GPH’s social media, including Facebook, YouTube, Twitter, Instagram, and Google Plus accounts. The study indicated that 84.75% of the 400-pilgrim sample found that GPH’s social media helped them know the services provided by GPH, which shows the significance of social media platforms to facilitate the performance of pilgrims’ rituals.

Many techniques can perform sentiment analysis (SA) for Twitter tweets [30]. Twitter tweets were tracked, and the sentiments of the tweeters were analyzed in various fields. For example, [31], [32] explored sports fans' feelings on Twitter during football matches and how their feelings were affected by the different events during the game. Bati [33] discussed the most prominent big data tools and techniques used in the SA field to collect, analyze, and represent data related to Twitter users' feelings and impressions and use these impressions as feedback to develop and improve the services provided. In another research, Zahrani et al. [34] collected and analyzed nearly five million Arabic and English tweets to monitor Twitter users' impressions of the Hajj of 1437 H (2016). According to specific keywords related to Hajj, the analyzed tweets were classified into two main categories, spatial and temporal. The analysis results showed a significant discrepancy in the impressions of the tweeters between positive, negative, and neutral due to many events that occurred during the pilgrimage, including spiritual and faith values that represent positive impressions. In contrast, the stampede in Mina, which caused some deaths, represents negative impressions.

The above-discussed studies show the importance of using social media to observe individuals and communities’ opinions, perceptions, and feelings, which can support recommendations to decision-makers and officials. Due to the changes caused by the Covid-19 that had a significant impact on the whole world, the Hajj season is not immune to these changes. Therefore, this paper mainly aims to investigate the social-related changes that occurred during the Hajj season of 1441 H within the community of Makkah.

3. Methodology

The research methodology used in this paper mainly consists of four steps summarized in Figure 1. The first step begins with reviewing the literature and previous studies related to tracking and analyzing the impressions and feelings of individuals in societies. Data collection is the second step that uses two methods to collect data from the Makkah community. The first was conducted through a questionnaire,
and the other by collecting tweets from the Twitter platform. Data analysis is the third step in which appropriate analysis procedures are applied. Finally, the last step includes the presentation of the results and discussion. An additional explanation for each step is given as follows:

3.1. Reviewing the literature

Reviewing previous research studies is an essential part of starting any research. It allows the researcher to investigate the last steps to handle similar problems and determine the research gaps. The previous section of this paper discussed some of the studies related to this paper’s topic.

3.2. Data collection

The data collection process consists of two main methods:

The first method is based on a 12-item questionnaire in which a random sample of 155 people from the Makkah community voluntarily participated. The questionnaire covers some information about the aspects of Hajj 1441 H, given in Table 1.

<table>
<thead>
<tr>
<th>Item</th>
<th>Demographics</th>
<th>Social aspect</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Nationality</td>
<td>The similarity in the attendance of Eid prayers for the Hajj season of 1441 H compared to previous years</td>
</tr>
<tr>
<td></td>
<td>Gender</td>
<td>The similarity of family gathering for breakfast on Eid for the Hajj season of 1441 H compared to previous years</td>
</tr>
<tr>
<td></td>
<td>Educational level</td>
<td>The continuation of the Eid celebration activities during the days of 11-13 Dhu al-Hijjah for the year of 1441 H</td>
</tr>
<tr>
<td></td>
<td>Age</td>
<td>The similarity of Eid celebration in the Hajj season of 1441 H compared to previous years</td>
</tr>
<tr>
<td></td>
<td></td>
<td>The degree of traffic congestion in Makkah during the Hajj season of the year 1441 H compared to previous years</td>
</tr>
<tr>
<td></td>
<td></td>
<td>The presence of volunteers in Makkah to serve pilgrims in the year 1441 H compared to previous years</td>
</tr>
<tr>
<td></td>
<td></td>
<td>The extent to which Makkah residents are interested in purchasing Eid-related goods in the Hajj season of 1441 H compared to previous years</td>
</tr>
<tr>
<td></td>
<td></td>
<td>The presence of street vendors during the Hajj season of 1441 H compared to previous years</td>
</tr>
</tbody>
</table>
The second method is based on collecting tweets from Twitter users within the holy city of Makkah using the Twitter application programming interface. The number of tweets used in this paper was about 45,000 tweets collected during the first 20 days of Dhul Al-Hijjah month, using a set of keywords including المساجد، النحر، عرفه، الحج، زحام، الإفطار، منى، الحرم، الحج، أضحية، شوارع، وصلاة العيد.

3.3. Data analysis

Data analysis is the third step of the methodology used, in which appropriate analysis methods were applied to analyze the data, and extract and visually present the results. Firstly, the questionnaire data were analyzed using essential statistical tools from which results and information were obtained and then visually represented using statistical charts.

Secondly, the collected tweets were analyzed using three main stages: preprocessing stage, text representation stage, and model architecture stage, as shown in Figure 2 [12], [13].

Figure 2. Components of the applied system

Stage 1. Data preprocessing:

The data preprocessing stage (also referred to as data normalization) is the primary stage in all SA methods. It includes a set of operations to process the source data in another format that allows it to be processed effectively. The procedures applied in this stage are identified according to the nature of the data and the application [18]. However, the preprocessing steps involved in this work are as follows [13]:

- Data cleaning: meaningless words and symbols are removed, including URLs, punctuation marks, special characters, and diacritics.
- Letter normalization: In this step, the different forms of letters are recognized and unified, such as the letters {ا, ا} are mapped to {ا}, {١} is mapped to {١} and {١} is mapped to {١}, as discussed in [23].
- Elongation Removal: Repetitive characters and letters are removed, such as {لا،لا،اجدد،ددد،لا،لا،اجدد،ددد}.

Stage 2. Text representation:

The skip-gram model is used to construct a new word embedding. The texts are represented as a 2D matrix $m \times n$, where $m$ is the number of words and $n$ is the embedding dimension, set to 300.

Stage 3. Model architecture

This stage presents two consecutive phases, Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) models. CNN model is used to extract features that are not expressed by the embeddings, whereas LSTM model detects the correlation between extracted
features. The output of LSTM is entered into softmax layer to generate the probability distribution of sentiment classes (i.e., positive, neutral, and negative). More details of the applied system components can be found in [12], [13].

3.4. Results and discussion

The result and discussion section is the last step of the research methodology used in this paper, where the obtained results are presented and discussed.

4. Results and discussion

This section provides the impressions of the Makkah community for the changes associated with the Hajj of 1441 H.

4.1. Questionnaire analysis results

The questionnaire data was collected for 155 participants during the Eid days in ten different locations in the holy city of Makkah. The results of the questionnaire were analyzed using essential statistical functions.

4.1.1. Demographic data

Figure 3 and Figure 4 show the nationality and gender of individuals who participated in the study questionnaire, respectively. Saudis represented 83.9% of the participants, whereas 16.1% were non-Saudis. The participation of females was more extensive with 62.6%, compared to males, whose participation did not exceed 37.4%.

The results show that the most significant proportion of the participants were university graduates with 60%, followed by post-university with 28.4%, as shown in Figure 5.

Figure 6 shows that 79.5% of the participants were between 20-and 50 years old, while only 1.3% were under 20-year-old.
4.1.2 A comparison between the attendance for the Eid prayer of Hajj 1441 H and previous years

Figure 7 illustrates the participants' perception of the similarity in the attendance for Eid prayers in Hajj of 1441 H compared to previous years. Results show that 2.6% strongly agree that the attendance of the Eid prayer of Hajj 1441 H is like previous years, whereas 25.8% strongly disagree. The results indicate an increase in the attendance to Eid prayers for the Hajj of the year 1441 H compared to the previous years. This increase is because most of Makkah's people were not busy serving the pilgrims that year due to adherence to Covid-19 precautionary measures.

![Figure 7. Participants' perception of the similarity in the attendance of Eid prayers in 1441 H compared to previous years](https://journals.kau.edu.sa/index.php/JENGSCI)

4.1.3 A comparison between the gathering of families for breakfast on Eid for the Hajj season of 1441 H and previous years

Figure 8 shows that most of the participants believe that the gathering of families for breakfast on Eid for the Hajj of 1441 H is different than previous years, with a total of 54.9%. These results denote a significant social-related change in the society of Makkah for the first time. Because of allowing a limited number of pilgrims to perform Hajj in 1441 H, families of Makkah society had the opportunity to have their Eid's breakfast together this year, which rarely happened in the previous years.

![Figure 8. Participants' perception of the similarity of family gathering for breakfast on Eid days in 1441 H compared to previous years](https://journals.kau.edu.sa/index.php/JENGSCI)

4.1.4 A comparison between the continuation of the Eid celebration in Hajj 1441

Figure 9 shows that the most significant percentage of participants (72.3%) did not extend the celebration of Eid in 1441 H to several days after the day of Eid. The limitation of the Eid celebration to only one day could be due to the continued spread of the COVID-19 pandemic.

![Figure 9. Participants' perception of the continuation of the Eid celebration in 1441 H compared to previous years](https://journals.kau.edu.sa/index.php/JENGSCI)
4.1.5 A comparison between the Eid celebration of Hajj 1441 H and previous years

Figure 10 shows the percentage of those who noticed the presence of celebration manifestations in Makkah during the Hajj of 1441 H. A total of 17.4% (for both: "yes definitely" and "yes, to some extent") believe that the celebration manifestations during Eid days of Hajj 1441 H are similar to previous years, while the most significant percentage of participants, 53.5% believe that there were no celebration manifestations. This result is consistent with the analysis of the previous question that the celebrations did not extend to the rest of Eid days.

4.1.6 A comparison between the degree of traffic congestion in Makkah during the Hajj season of 1441 H and previous years

The results indicate that 83.2% of the participants found that the traffic congestion was lower than in previous years, as shown in Figure 11. In contrast, about 3.2% of the participants believed that the degree of traffic congestion was the same or greater than the previous years.

4.1.7 A comparison between the presence of volunteers in Makkah to serve pilgrims during the Hajj of 1441 H and previous years

The results illustrate that 47.1% of the participants believed that the presence of volunteers is less in Hajj 1441 H compared to previous years, as shown in Figure 12. The results show that 49% of the participants are unaware of the presence of volunteers who serve pilgrims, which is due to the adherence to the instructions of the Ministry of Health to stay at home.
4.1.8 A comparison between the interest of Makkah residents to purchase Eid-related goods in the Hajj season of 1441 H and previous years

Figure 13 shows that 45% of participants believed that the interest in purchasing Eid-related goods in the Hajj season of 1441 H is less compared to previous years. This observation results from the fear of many people from being in public places due to the outbreak of Covid-19. The results also show that 47% were unaware of the difference.

Figure 13. Participants’ perception of the interest to purchase Eid-related goods in the Hajj season of 1441 H compared to previous years

4.1.9 A comparison between the presence of street vendors during the Hajj season of 1441 H and previous years

The results in Figure 14 demonstrate that the most significant percentage of the participants, 49%, are unaware of whether there are street vendors. However, 46.5% of the participants indicated that the presence of street vendors was lower than in previous years.

Figure 14. Participants’ perception of the presence of street vendors during the Hajj season of 1441 H compared to previous years

4.2 Tweet’s analysis results

Several keywords were used to collect tweets from the Twitter platform. The collected tweets are then analyzed, and SA values are extracted according to the method used in [12], [13]. It should be noted that the average SA score for all tweets is between 1 and 5, where 1 represents the minimum sentiment value of "not quite satisfied" and 5 represents the maximum sentiment value of "extremely satisfied."

4.2.1 Analyzing tweets using specific keywords related to Hajj

4.2.1.1 Analysis of the tweets of Makkah people for the keyword "عرفة" ("Arafa").

During the Hajj days of 8-13 Dhu Al-Hijjah 1441 H, the number of tweets collected that include the keyword "عرفة" within the city of Makkah was 977, 59% of which were recorded on the 9th of Dhu Al-Hijjah, as shown in Figure 15. Results show that the SA rate ranged between 4.02 to 4.75, and the overall average was 4.38 for all days, reflecting the happiness of the tweeters, especially on day 9.

Figure 16 shows the sentiment classes rate of tweets for the keyword "عرفة," which were categorized into three classes: positive, neutral, and negative. The results show that positive tweets recorded the highest rate in all Hajj days, while negative tweets were the lowest.
4.2.1.2 Analysis of the tweets of Makkah people for the keyword "الحجاج" ("Al-Hujaj")

For the keyword "الحجاج," 169 tweets were collected during the Hajj days (8-13 Dhu Al-Hijjah), 55.6% of which were recorded on the 8th of Dhul Al-Hijjah, as shown in Figure 17. The results show that the SA rate ranged between 3.4 and 4.8, with a total average of 4.15, indicating the happiness rate and satisfaction of the tweeters on most Hajj days.

Figure 18 shows that the rate of positive tweets was the highest on all Hajj days except for the tenth day, where the neutral tweets were the largest due to the emergence of some news related to the movement and transportation of pilgrims.
3.2.1.3. Analysis of the tweets of Makkah people for the keyword "العـيد" ("Eid")

Regarding the keyword "العـيد," the number of tweets collected was 12,165 recorded during the days of 9-14 Dhul-Al-Hijjah. The largest percentage of tweets was recorded on day 9 of Dhul Al-Hijjah with 42%, as shown in Figure 19. The results show the SA rate was between 3.69 and 4.33, with a total average of 3.97 for all Hajj days, indicating high satisfaction and happiness of the tweeters in the celebration of Eid.

Figure 20 shows the rate of tweets categorized as positive, neutral, and negative tweets for the keyword "Eid". The figure shows that positive tweets were the most common on all days of Hajj except for days 12 and 13, where neutral tweets were the most.
3.2.1.4. Analysis of the tweets of Makkah people for the keyword "الحج" ("Hajj")

Figure 21 shows the percentage rates of 1,442 tweets collected for the days of 1-13 Dhul Al-Hijjah, including the keyword "الحج". The recorded results show that days 8 and 9 Dhul Al-Hijjah were the largest, with 31.5% and 23.2%, respectively. The results show that the SA rate before day 8 (i.e., before starting Hajj rituals) was between neutrality and satisfaction degrees. In contrast, the SA rate changed on day 8 from less than satisfaction degree (3.59) to higher than satisfaction degree (4.28). In other words, the start of the Hajj rituals shows a positive impact on the emotions and feelings of the tweeters, which reflects their happiness.
Figure 21. SA rate for tweets that include the keyword "Hajj"

Figure 22. Sentiment classes for tweets that include the keyword "Hajj"

Figure 22 presents the sentiment classes rate of tweets that include the keyword "Hajj" for the days 8, 9, and 10 Dhul Al-Hijjah. The results show that the positive tweets achieve the highest rates on all days, which indicates the happiness and positive impression of the Makkah people for the coming hajj days, especially on the days 8 and 9 Dhu Al-Hijjah.

Moreover, most tweets under the neutral sentiment category include news related to Hajj events, instructions, and regulations. Most of the negative impressions of the tweets include feeling sad about not being able to see roads, shops, hotels, restaurants, and other places as crowded with pilgrims as they were in the previous years.

3.2.2. Analysis of tweet types

Figure 23 shows the type of tweets in the original tweet, retweet, or replied tweet. The most significant proportion was the response to tweets among users, with 43% of the total tweets, which indicates the large volume of participation and interaction by replying to original tweets. The percentage of original tweets and retweets was 36.2% and 20.5%, respectively.
3.2.3. Analysis of tweet source

Figure 24 presents the primary sources through which the tweets were posted. Most tweets were posted through iPhone devices with 77.3%, followed by devices running on Android systems with 19.3%, and then Web Apps with 2.2%.

![Figure 23. Tweet types](image1)

![Figure 24. Tweet sources](image2)

5 Conclusion and future work

In this paper, the most prominent social aspects of the Makkah community during the Hajj season of 1441 H were studied and analyzed, focusing on the social-related changes in the lifestyle of the Makkah community under the circumstances of the COVID-19 pandemic compared to previous Hajj seasons. Two methods were used to collect data from the Makkah community using a questionnaire and the Twitter platform. The results show an increase in the Eid prayer of Hajj 1441 H attendance compared to previous years, which is one of the new social manifestations that does not usually occur.

The results also reveal that families are more interested in tacking Eid of Hajj breakfast together in 1441 H compared to previous years. In addition, this paper noticed some other observations that emerged in Hajj of 1441 H due to COVID-19, including Eid celebration manifestations, traffic congestion, presence of volunteers, interest in shopping, and the presence of street vendors.

The analysis results for 14,753 tweets showed positive results for sentiment analysis, where the overall rate of sentiment analysis ranged between satisfaction and more than satisfied. Furthermore, happiness was noticeable in the...
tweets during all the days of Hajj, especially after the start of the Hijj rituals. On the other hand, some tweets contained feelings of longing and sadness for not seeing the places of Makkah crowded with pilgrims, compared to previous years, which were classified as negative tweets.

Future work may include using data mining tools to study the impact of COVID-19 on human lifestyle in the cities of Makkah and Medina. In addition, the study will continue to collect and analyze tweets from the Makkah and Medina communities and conduct comparative studies before and after the end of COVID-19.

References


الملخص

على مر التاريخ، استمر مجتمع مكة بفضل قصارى جهده لخدمة المسلمين الذين يأتون لأداء مناسك الحج (العمرة)، الركن الخامس من أركان الإسلام. إلا أن حج عام 1441 هـ (2020 م) كان مختلفًا بسبب تفشي فيروس كورونا الجديد (كوفيد-19). وقد أحدث هذا الأثر تغييرات عديدة في العديد من جانب الحياة البشرية في مختلف المجتمعات، بما في ذلك مجتمع مكة المكرمة. بينما كان أهل مكة يستعدون لتقييم ما اعتادوا القيام به كل عام لدعم وتسهيل الخدمات للحجاج، كان عليهم اتباع تعليمات وزارة الصحة لتنفيذ العديد من الإجراءات الاحترازية والوقائية لتجنب الإصابة بفيروس 19؛ وأهمها البقاء في المنزل. تهدف هذه الدراسة بشكل أساسي إلى مراقبة وتحليل التغييرات الاجتماعية التي قد تكون حاضرة أو غائبة لأول مرة في مجتمع مكة المكرمة أثناء حج 1441 هـ. تم استخدام طريقتين رئيستين: تحليل المشاعر المعتمد على الذكاء الاصطناعي لعينة من التغريدات. وكذلك استبانة لعينة عشوائية من سكان مكة المكرمة.