IJICTR
International Journal of Information & Communication Technology Research

Volume 14- Number 2 – 2022 (23 -31)

Large-Scale Twitter Mining for Extracting the Psychological Impacts of COVID-19

Hamed Vahdat-Nejad*
Perlab, Faculty of Electrical and Computer Engineering
University of Birjand
Birjand, Iran
vahdatnejad@birjand.ac.ir

Faezeh Azizi
Perlab, Faculty of Electrical and Computer Engineering
University of Birjand
Birjand, Iran
faezeh.azizi1995@birjand.ac.ir

Mahdi Hajiabadi
Perlab, Faculty of Electrical and Computer Engineering
University of Birjand
Birjand, Iran
mahdihajiabadi@birjand.ac.ir

Fatemeh Salmani
Perlab, Faculty of Electrical and Computer Engineering
University of Birjand
Birjand, Iran
salmani_fatemeh98@birjand.ac.ir

Sajedeh Abbasi
Perlab, Faculty of Electrical and Computer Engineering
University of Birjand
Birjand, Iran
sajedeh_abbasi@birjand.ac.ir

Mohadese Jamalian
Perlab, Faculty of Electrical and Computer Engineering
University of Birjand
Birjand, Iran
mohadesejamalian@birjand.ac.ir

Reyhane Mosafer
Perlab, Faculty of Electrical and Computer Engineering
University of Birjand
Birjand, Iran
reyhane.mosafer@birjand.ac.ir

Hamideh Hajiabadi
Department of Computer Engineering
Birjand University of Technology
Birjand, Iran
hajiabadi@birjandut.ac.ir

Wathiq Mansoor
Department of Electrical Engineering
University of Dubai
Dubai, UAE
wmansoor@ud.ac.ae

Received: 14 December 2022 – Revised: 19 May 2022 - Accepted: 26 June 2022

Abstract—**The outbreak of the COVID-19 in 2020 and lack of an effective cure caused psychological problems among humans. This has been reflected widely on social media. Analyzing a large number of English tweets posted in the early stages of the pandemic, this paper addresses three psychological parameters: fear, hope, and depression. The main issue is the extraction of the related tweets with each of these parameters. To this end, three lexicons are proposed for these psychological parameters to extract the tweets through content analysis. A lexicon-based method is then used with GEO Names (i.e. a geographical database) to label tweets with country tags. Fear, hope, and depression trends are then extracted for the entire world and 30 countries. According to the analysis of results, there is a high correlation between the frequency of tweets and the official daily statistics of active cases in many countries. Moreover, fear tweets dominate hope tweets in most countries, something which shows the worldwide fear in the early months of the pandemic. Ultimately, the diagrams of many countries demonstrate unusual spikes caused by the dissemination of specific news and announcements.**

Keywords: natural language processing; emotion analysis; knowledge extraction; data mining.

Article type: Research Article

Publisher: ICT Research Institute

* Corresponding Author

I. INTRODUCTION

Social media such as Twitter and Facebook are known as rich sources of user opinions. With the occurrence of important events around the world, users of social media share a huge number of posts and reviews, which sometimes contain noteworthy hidden information. In addition to reportage and marketing, Twitter is a social medium also used for opinion mining and identification of social behavior shown by users. For this purpose, natural language processing is employed to analyze user opinions on various areas such as politics [1], education [2], health [3], natural disasters [4], and economy [5] to extract implicit information such as the prediction of trends in the stock market [6]. A few studies have also been conducted on different epidemics, such as the flu [7], MERS [8], and Ebola [9], reporting significant results such as the disease prediction model [8].

Since 2020, the novel coronavirus has been among the hottest topics on social media, about which millions of tweets are posted on a daily basis [10]. Emotion analysis can be employed to perceive people's emotions regarding different parameters such as hope, fear, and depression, which play a central role in their moods. In fact, quarantine and lockdown regulations have been set in cities and countries since the outbreak of the COVID-19 pandemic; as a result, welfare and economic conditions have become harsher than ever before and infused societies with depression [11] and fear [12]. On the one hand, the increasing number of active cases and deaths from the COVID-19 decreased life expectancy [13]. On the other hand, hopefulness has been escalating among people in some countries where political, economic, and hygienic actions were taken to contain the COVID-19.

During this pandemic, different studies have been conducted on the static analysis of emotions of tweets in terms of psychological parameters [14, 15]. This paper conducts a dynamic analysis of tweets regarding hope, fear, and depression to identify the trend in each of these psychological parameters in different countries on the onset of the COVID-19 pandemic. For this purpose, more than two million relevant tweets are analyzed within 14 weeks (March 23 to June 23, 2020) with respect to hope, fear, and depression. A comprehensive lexicon is first created for fear and hope by using NRC [1] and GI [2] lexicons as well as the WORDNET[3] ontology (*i.e.*, a complete reference of terminologies). Since there are no appropriate reference lexicons for depression, an in-house lexicon is created for this parameter. The relevant tweets of every parameter are then extracted by leveraging the proposed lexicons. The tweets are then labeled with geotags by collecting a geo lexicon (including names of countries, provinces/states, cities, *etc.*) using Geo Names[4], which is a comprehensive database of names of locations such as countries, provinces/states, and cities. In fact, the content analysis of a geographical location through a lexicon can help accurately retrieve the location of interest in tweets. Finally, the trend in every parameter over time is drawn for the whole world and 30 countries. The analysis of trend diagrams yields interesting results.

This paper extends our previous work [16], in which we extracted the major topics discussed by Twitter users regarding COVID-19. The contributions of this paper are as follows:

- A large dataset of more than two million tweets regarding the novel coronavirus is employed to conduct an emotion analysis. The larger the dataset in data mining applications, the more reliable the extracted knowledge.
- The geographical location of tweets is determined through content analysis using a geo lexicon.
- The relevant tweets are extracted and analyzed based on separate lexicons proposed for hope, fear, and depression.

The rest of this paper consists of different sections. A literature review is presented in Section 2. Section 3 introduces the proposed method, which is implemented in Section 4. Section 5 includes the discussion, and finally, Section 6 presents the research results and future studies.

II. RELATED WORK

The emotion analysis of user comments on diseases and epidemics can be employed to measure society's feelings in the face of disease. In this regard, 12101 tweets posted in March 2015 were analyzed to evaluate the public opinion on the Ebola virus disease. The tweets were then classified into six categories, including anger, disgust, happiness, sadness, surprise, and fear [17]. Moreover, South Korean people's emotions were categorized into seven classes (*i.e.* neutral, happiness, sadness, anger, disgust, fear, and surprise) to analyze their reactions to the spread of the 2015 Middle East respiratory syndrome (MERS). The results indicate an increase in the number of comments of anger over time, something which shows people's anger with the crisis [17]. Furthermore, a few tweets posted in 2018 were extracted and classified through the emotion analysis into three categories (*i.e.*, supporting vaccination, rejecting vaccination, and neutral) to evaluate differences in various emotions towards vaccination over time [18].

Since the beginning of 2020, social media users have shared their experiences and observations regarding the outbreak of the COVID-19 on different platforms, especially on Twitter. Some studies have been conducted to analyze these comments. To evaluate the public opinion on the enforcement of social distancing regulations in US, 259529 tweets were collected from January 23 to March 23, 2020. After that, social distancing facets (including negative emotion, implementation, social disruption, positive emotion, adaptation, and purpose) have been extracted from the tweets for emotion analysis [19]. Similarly, 30000 English tweets were collected from January 22 to April 15, 2020, to evaluate the impact of COVID-19

---

[1] https://saifmohammad.com/WebPages/NRC-Emotion-Lexicon.htm
[2] http://www.wjh.harvard.edu/~inquirer/homecat.htm
[3] https://wordnet.princeton.edu/
[4] https://www.geonames.org/

on the mental health of Twitter users worldwide. The NRC word-emotion lexicon has then been employed to label the tweets with emotion tags and determine their sentiment scores based on different parameters of emotions such as positive emotion, negative emotion, joy, sadness, anger, fear, trust, disgust, anticipation, and surprise [20]. Furthermore, 410643 tweets about the COVID-19 were collected from India from March 22 to April 21, 2020, to extract different emotions using the NRC lexicon [21].

Similarly, 30000 tweets about the COVID-19 were analyzed separately for the USA, the UK, Spain, Sweden, Italy, and Germany in April 2020. The emotion proportion of tweets was then determined for each country through emotion analysis [15]. Finally, 16138 tweets have been analyzed to evaluate behavioral changes of Twitter users during the COVID-19 pandemic within three different intervals (February 5–11, May 21–27, and June 15–21). The emotion analysis has then been conducted to extract such emotions as confidence, anger, analytics, sadness, joy, fear, and tentativeness. The geotags of tweets have then been employed to draw the distribution of tweet emotions for Canada, France, India, Italy, Spain, the UK, and the USA for the three designated intervals. According to the results, tweets of sadness have the highest percentage comparing with other emotions over the entire intervals [14].

Following the previous studies, this paper conducts a dynamic analysis of emotions over time to determine the trends in fear, depression, and hope. Unlike the previous studies, geotags are not utilized in this paper in order to improve accuracy, for a tweet might be about a country that does not match the user's geographical location (for instance, a user based in China might tweet about the US). To this end, a geo content analysis method is proposed for location extraction based on the locations mentioned in tweets. Finally, since large datasets make results closer to reality in NLP applications, contrary to similar studies, a large dataset of two million tweets is considered.

III. TWEET PROCESSING

This study aims to extract user emotions regarding the COVID-19 during the early months of the pandemic. For this purpose, a large number of tweets about the COVID-19 have been selected and emotionally processed based on fear, hope, and depression.

A. Data set

COVID-19 was identified as a pandemic by the World Health Organization (WHO) in the middle of March 2020 [22]. Since this study aims to analyze worldwide people's emotions during the first three months of the pandemic, the research interval has been from March 23 to June 23, 2020. For every week, 150000 tweets have been selected randomly based on keywords including "corona", "coronavirus", "COVID", "pandemic", "SARS-CoV-2", and "COVID-19".

B. Data set analysis

The first processing step is identifying the tweets related to fear, hope, and depression. For this purpose, a content analysis method is proposed based on the development of lexicons for the abovementioned emotions. Each of these lexicons is developed differently from another. Created through the NRC lexicon[5], the fear lexicon contains 1300 expressions such as "concerned", "blackmail", and "bloody endangered". Developed through GI [6] (General Inquirer), the hope lexicon contains 500 expressions such as "positivistic", "joy", and "gladden". Since there are no appropriate reference lexicons that can be employed to develop a depression lexicon, an in-house lexicon is created to contain 100 expressions such as "dimple", "dark", "agony". For this purpose, Wordnet[7] ontology has been used to complete and further develop the keywords of each lexicon. Using Wordnet, we get all the synonymous words with the primitive lexicons. Unrelated keywords are then filtered and removed.

The relevant tweets are then extracted using the designated lexicons and implementing a GATE [23] pipeline. The frequency of relevant tweets for each parameter indicates the prevalence of that emotion parameter. Variations in the frequency of tweets can also show the change in trends of emotion parameters over time.

After extracting the datasets related to fear, hope, and depression, the content location of each tweet should be determined. For this purpose, a list of designated countries with their states and cities is collected through Geo Names[8], which is a geographical database containing more than 25 million names of locations and geographical contents such as names of countries, states/provinces, and cities. The prevalence of COVID-19 is considered in different countries to select 30 countries[9] with the highest formal infection statistics for analysis. After that, the designated names of states/provinces and cities are codified on a list containing nearly 7000 terms. The words of every tweet are then compared with this list. If the name of a geographical location exists in a tweet, the corresponding country is identified and attributed to the tweet as a feature.

Figs. 1 to 3 show the word cloud related to the tweets of fear, hope, and depression, respectively. They highlight the repetitive words in the tweets of each emotion. In the word cloud of fear, keywords such as "fight corona", "black live", "Airborne virus", and "dangerous people" are frequent. In the word cloud of hope, keywords such as "lives matter" and "please wear" have the most repetitions in tweets. Finally, in the word cloud of depression, keywords such as "let sink", "hit million", and "stop thinking" are more frequent.

[5] The NRC Emotion Lexicon is a list of English words and their associations with eight basic emotions (anger, fear, anticipation, trust, surprise, sadness, joy, and disgust) and two sentiments (negative and positive).
[6] http://www.wjh.harvard.edu/~inquirer/homecat.htm
[7] https://wordnet.princeton.edu/
[8] http://download.geonames.org/export/dump
[9] Australia, Belarus, Belgium, Brazil, Canada, Chile, China, Ecuador, France, Germany, India, Iran, Ireland, Italy, Japan, Mexico, Netherlands, Pakistan, Peru, Qatar, Russia, Singapore, South Korea, Spain, Sweden, Switzerland, Turkey, UAE, UK, USA

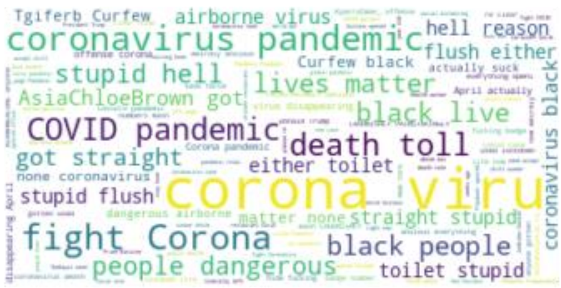


Fig. 1. Word cloud related to the fear.

Fig. 2. Word cloud related to the hope

Fig. 3. Word cloud related to the depression

IV. EXPERIMENT

The initial dataset contains more than two million tweets regarding the novel coronavirus from March 23 to June 23, 2020. In fact, 150000 tweets have been gathered in every week of the designated interval. We have analyzed the text of the tweets in GATE, which is a "platform for developing and deploying software components that process human language" [24].

GATE has a few natural language processing components, one of which is tagging. To extract the tweets related to every emotional parameter (*e.g.* hope), the relevant list is used as a Gazetteer list in a GATE pipeline. The tweets are then compared with the list. Similarly, the location list is also entered into the GATE pipeline. After each comparison, the corresponding country's name is selected as the tweet feature.

After that, the emotion parameters are analyzed for the entire world and the designated countries. Fig. 4 reports the frequency of tweets regarding every emotion parameter (*i.e.*, hope, fear, and depression) for the entire world on a weekly basis from March 23 to June 23, 2020. Moreover, the official statistics of active cases of the COVID-19 are obtained from John Hopkins University and Medicine Coronavirus Resource Center (https://coronavirus.jhu.edu/map.html).

According to Fig. 4, there are more tweets of fear than those of hope and depression in all 14 weeks of analysis. This indicates the dominance of fear as well as people's ever-increasing panic about the prevalence of the COVID-19. As a result, fear affected people's thoughts and emotions more than anything else during the initial months of the pandemic. The diagrams analysis of fear and hope indicates that hope declines when fear escalates. In other words, the instant rise of fear among people diminishes their hopefulness and vice versa. In fact, in most of the weeks, these two curves are symmetric. Ultimately, the fear curve shows an upward trend during the investigated period in which the official statistics of active cases of the COVID-19 is also on the rise.

Fig. 5 demonstrates the frequency of tweets about hope, fear, and depression in the 30 countries with the highest statistics of active cases. According to the diagrams of different countries, there is a correlation between tweet frequency curve trends and the official statistics of active cases in many countries such as Germany, Ecuador, Italy, Spain, Brazil, Belgium, Iran, Qatar, Singapore, and Turkey. At the same time, China's diagram shows a downward trend in the official statistics of active cases from the sixth week to the twelfth week, whereas the tweet frequency curve reached its peak. The analyses indicate that China was among the most discussed countries among worldwide users in that interval. Moreover, users posted most of their tweets in India and Pakistan weeks before the global outbreak of the COVID-19, the reasons for which can be people's concerns about the prevalence of this disease in populous countries.

According to the diagrams of several countries, there are nearly zero depression tweets; thus, the tweets related to the COVID-19 are usually classified as either fear or hope tweets. Likewise, there is a significant correlation between the official statistics of active cases and the number of fear tweets in various countries such as Brazil, Spain, Peru, and the Netherlands. In other words, the number of fear tweets escalates as the number of active cases increases, whereas the number of these tweets declines as the number of active cases decreases. In most countries, there are more fear tweets than hope tweets, something which shows that people panicked in the early months of the COVID-19 pandemic. At the same time, there are more hope tweets than fear tweets in a few countries such as Germany and Singapore in most weeks. This can be due to useful actions taken by governments for political, economic, and hygienic purposes.

Ultimately, the number of tweets escalates abnormally in a few other countries during one or some specific weeks due to the occurrence of specific events or spread of specific news such as people's dissatisfaction with socioeconomic policies of governments, the COVID-19 containment management, passiveness, and dishonesty of officials, and escalated tyranny and suppression. In addition, border lockdown in Qatar and illegal immigrants or foreign travelers in Mexico and Britain made people worried about the rise in the COVID-19 spread. Moreover, users have posted many tweets about earthquakes in Mexico, the spread of the flu in the US, and fires and riots in Australia. People have also been worried by the dissemination of news on unemployment, declined monthly income, poverty of families, and bad economic conditions in Spain and Britain. Finally, many tweets have been posted about racist attacks in the UAE and Canada as well as the large number of inmates in Russia and Turkey.

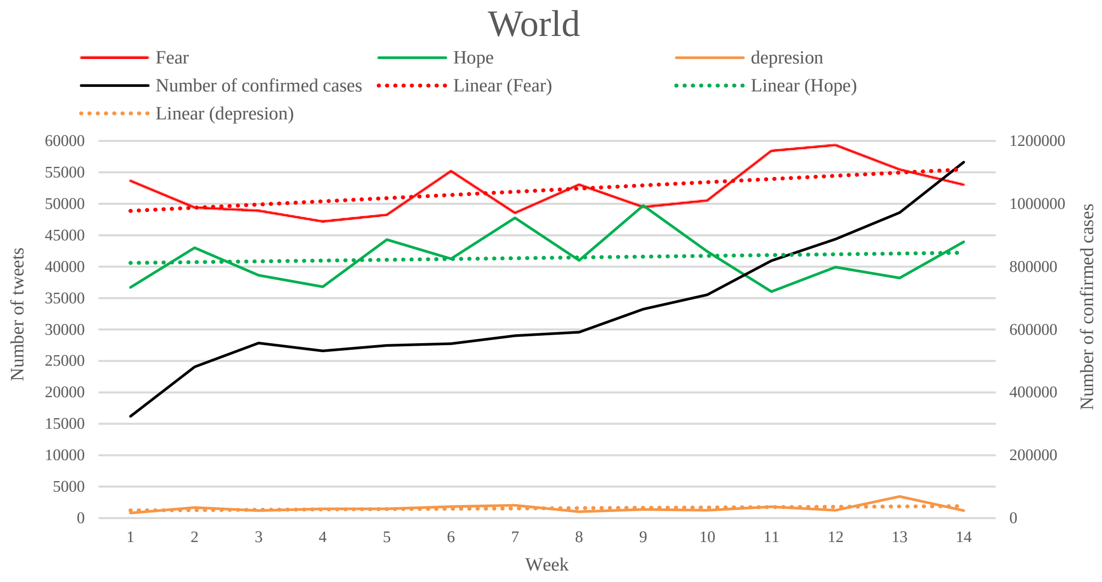


Fig. 4. Weekly frequency of tweets of fear, hope, and depression as well as the official statistics of active cases of the COVID-19 worldwide

The unprecedented rise in depression tweets is only observed in the UAE, where most of the tweets are about the discriminatory behavior of Indian immigrants towards Muslims.

V. Discussion

This paper examines the three emotions of fear, hope, and depression in users' tweets about COVID-19 on Twitter. However, it should be noted that extracting human emotions from text is complicated [25]. For example, the tweet "I hope you licked up corona virus" contains the word hope but has a negative connotation; therefore, over-generalization should be prevented.

We see some motivating remarks that could pave the way for future research to investigate further the psychological impacts of people on the COVID-19 epidemic. In this regard, the immediate increase in fear among people (Fig. 4) can be the increase in the number of confirmed cases and the number of deaths.

Another key finding of this paper is that users' feelings about different countries are not the same. For example, there is a greater sense of hope in East Asian countries due to the timely suppression of the COVID-19 virus. In contrast, positive news such as humanitarian aid in the Netherlands, Switzerland, Ireland, Germany, and Canada, political, economic, and hygienic actions of governments to contain the COVID-19 and mitigate economic pressure on families in Singapore, South Korea, and Germany, and foreign border lockdown in Mexico, the US, and Canada have resulted in an occasional increase in the number of hope tweets.

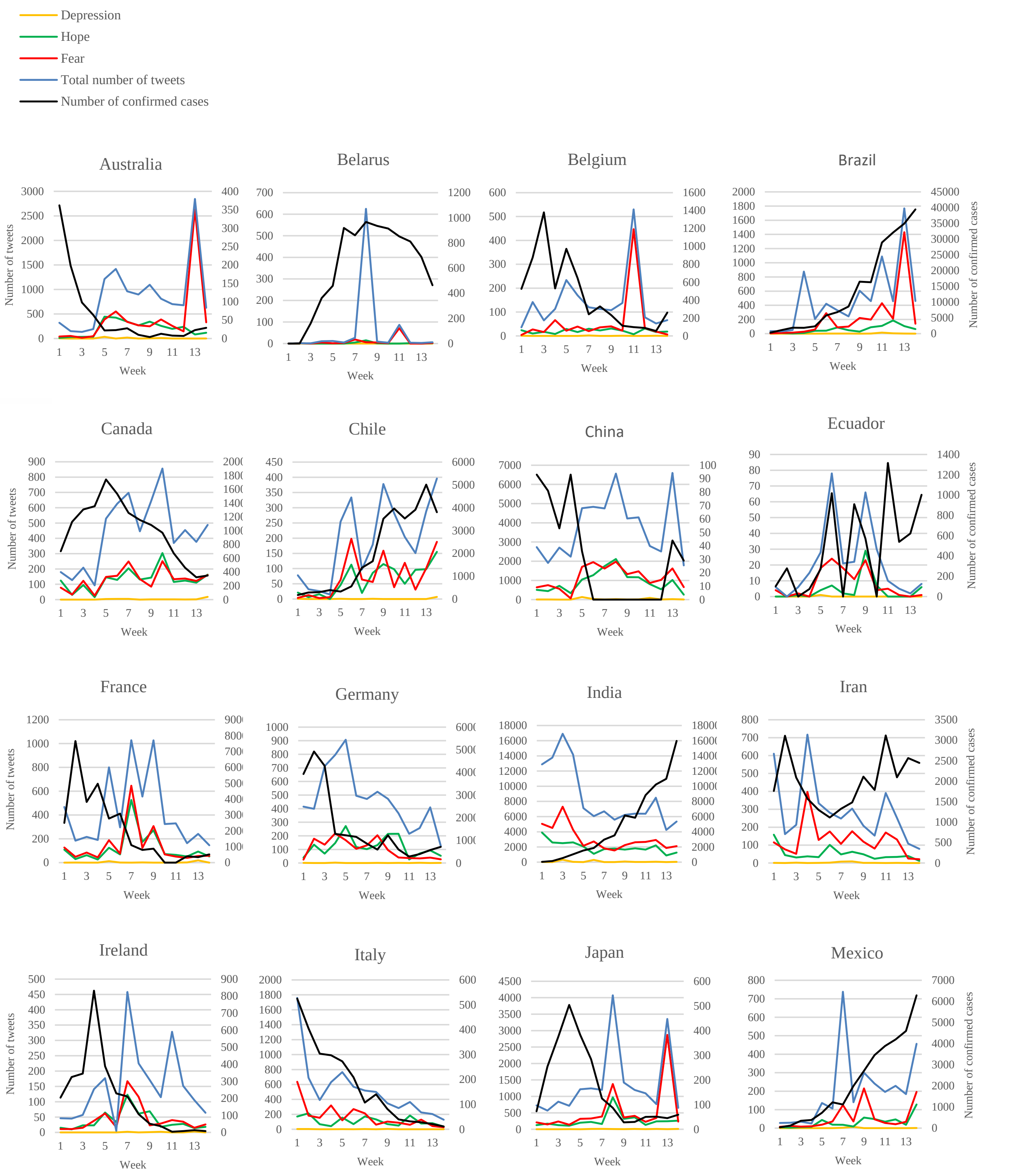
Generalizing these cases to the general population of the countries mentioned is not completely accurate. In fact, we should see how much the cultural, welfare, political, social, healthcare, educational, and economic parameters might affect the emotional changes during the COVID-19 outbreak?".

Our study showed that while different emotions are common in COVID-19 tweets in different countries, the number of confirmed cases is highly correlated with the number of tweets containing the fear emotion in most countries. (Fig. 5).

VI. Conclusion

This paper has analyzed the tweets about the COVD-19 in the early stages of the pandemic based on three psychological parameters, *i.e.*, fear, hope, and depression. Separate lexicons have been proposed for these parameters by using NRC and GI lexicons as well as the WORDNET ontology. The relevant tweets of each parameter were then extracted by implementing the respective lexicons in the GATE pipeline. After that, the Geo Names database (*i.e.*, a comprehensive database of names of geographical locations) has been employed to collect a comprehensive lexicon of locations. The tweets were then labeled with geotags through a content analysis method. Finally, each parameter trend has been extracted for the entire world and the 30 countries having the highest statistics of active COVID-19 cases. According to the analysis results, fear was the dominant parameter in most countries. In addition, hope was more prevalent in Eastern Asian countries, which could be due to the timely containment of the novel coronavirus and the resolution of this turbulent crisis. In the end, there was a correlation between the frequency of tweets and the official statistics of active cases in most countries.

Since a considerable number of tweets are non-English, a language-independent method can be proposed to reach a more comprehensive and more accurate analysis of the public opinions. Given the inherent differences between various languages around the world, future studies can consider any of the non-English languages such as French, Chinese, Persian, and Arabic for analysis and evaluation. Moreover, the tweets about the COVID-19 can be classified into the economy, education, tourism, psychology, and other areas to analyze the tweets of each domain based on their specific parameters.



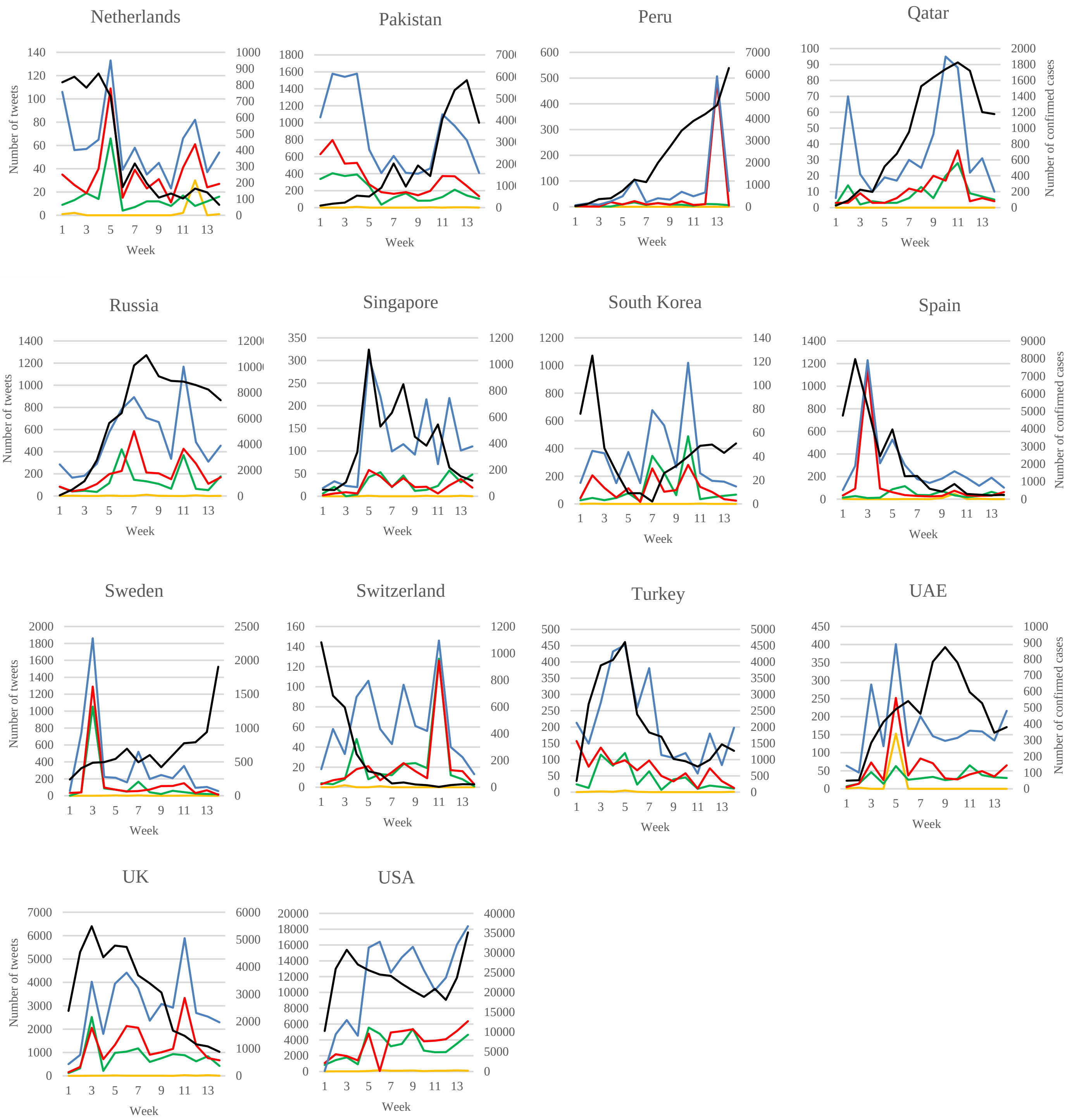


Fig. 5. Energy Frequency of tweets of hope, fear, and depression for the 30 designated countries

References

[1] L. Belcastro, R. Cantini, F. Marozzo, D. Talia, and P. Trunfio, "Learning Political Polarization on Social Media Using Neural Networks," *IEEE Access*, vol. 8, pp. 47177-47187, 2020.

[2] G. Marinoni, H. Van't Land, and T. Jensen, "The impact of Covid-19 on higher education around the world," *IAU Global Survey Report*, pp. 1-50, 2020.

[3] D. Mahata, J. Friedrichs, R. R. Shah, and J. Jiang, "Detecting personal intake of medicine from twitter," *IEEE Intelligent Systems*, vol. 33, no. 4, pp. 87-95, 2018.

[4] T. Dereli, N. Eligüzel, and C. Çetinkaya, "Content analyses of the international federation of red cross and red crescent societies (ifrc) based on machine learning techniques through twitter," Natural Hazards, pp. 1-21, 2021.

[5] J. Bollen, H. Mao, and X. Zeng, "Twitter mood predicts the stock market," *Journal of computational science*, vol. 2, no. 1, pp. 1-8, 2011.

[6] D. Valle-Cruz, V. Fernandez-Cortez, A. López-Chau, and R. Sandoval-Almazán, "Does Twitter Affect Stock Market Decisions? Financial Sentiment Analysis During Pandemics: A Comparative Study of the H1N1 and the COVID-19 Periods," *Cognitive computation*, pp. 1-16, 2021.

[7] B. Alkouz, Z. Al Aghbari, and J. H. Abawajy, "Tweetluenza: Predicting flu trends from twitter data," *Big Data Mining and Analytics*, vol. 2, no. 4, pp. 248-273, 2019.

[8] D.-H. Choi, W. Yoo, G.-Y. Noh, and K. Park, "The impact of social media on risk perceptions during the MERS outbreak in South Korea," *Computers in Human Behavior*, vol. 72, pp. 422-431, 2017.

[9] C. Priest and D. Groves, "Tweeting about Ebola: Analysis of Tweets from Africa, Europe and the United States During Two Months of the 2019 Ebola Virus Disease (EVD) Epidemic in the Democratic Republic of the Congo," in *International Conference on Information and Communication Technologies for Disaster Management*, Paris, France, 2019: IEEE, pp. 1-2.

[10] E. Chen, K. Lerman, and E. Ferrara, "Tracking Social Media Discourse About the COVID-19 Pandemic: Development of a Public Coronavirus Twitter Data Set," *JMIR Public Health and Surveillance*, vol. 6, no. 2, pp. 1-9, 2020.

[11] M. Z. Ahmed, O. Ahmed, Z. Aibao, S. Hanbin, L. Siyu, and et al, "Epidemic of COVID-19 in China and associated psychological problems," *Asian journal of psychiatry*, vol. 51, pp. 1-7, 2020.

[12] D. Doshi, P. Karunakar, J. R. Sukhabogi, J. S. Prasanna, and S. V. Mahajan, "Assessing coronavirus fear in Indian population using the fear of COVID-19 scale," *International Journal of Mental Health and Addiction*, pp. 1-9, 2020.

[13] S. Trias-Llimós and U. Bilal, "Impact of the COVID-19 pandemic on life expectancy in Madrid (Spain)," *Journal of Public Health*, vol. 42, no. 3, pp. 635-636, 2020.

[14] S. Kaur, P. Kaul, and P. M. Zadeh, "Monitoring the Dynamics of Emotions during COVID-19 Using Twitter Data," *Procedia Computer Science*, vol. 177, pp. 423-430, 2020.

[15] G. Matošević and V. Bevanda, "Sentiment analysis of tweets about COVID-19 disease during pandemic," in *43rd International Convention on Information, Communication and Electronic Technology*, Opatija, Croatia, 2020: IEEE, pp. 1290-1295.

[16] F. Azizi, H. Vahdat-Nejad, H. Hajiabadi, and M. H. Khosravi, "Extracting Major Topics of COVID-19 Related Tweets," presented at the Eleventh International Conference on Computer and Knowledge Engineering, Mashhad, 2021.

[17] H. J. Do, C.-G. Lim, Y. J. Kim, and H.-J. Choi, "Analyzing emotions in twitter during a crisis: A case study of the 2015 Middle East Respiratory Syndrome outbreak in Korea," in *International conference on big data and smart computing*, Hong Kong, China, 2016: IEEE, pp. 415-418.

[18] R. Mahajan, W. Romine, M. Miller, and T. Banerjee, "Analyzing Public Outlook towards Vaccination using Twitter," in *International Conference on Big Data*, Los Angeles, USA, 2019: IEEE, pp. 2763-2772.

[19] J. Kwon, C. Grady, J. T. Feliciano, and S. J. Fodeh, "Defining facets of social distancing during the covid-19 pandemic: Twitter analysis," *Journal of Biomedical Informatics*, vol. 111, pp. 1-7, 2020.

[20] A. Mathur, P. Kubde, and S. Vaidya, "Emotional Analysis using Twitter Data during Pandemic Situation: COVID-19," in *5th International Conference on Communication and Electronics Systems*, Coimbatore, India, 2020: IEEE, pp. 845-848.

[21] S. Das and A. Dutta, "Characterizing public emotions and sentiments in COVID-19 environment: A case study of India," *Journal of Human Behavior in the Social Environment*, pp. 1-14, 2020.

[22] E. Gharavi, N. Nazemi, and F. Dadgostari, "Early outbreak detection for proactive crisis management using twitter data: Covid-19 a case study in the us," *arXiv preprint arXiv:2005.00475*, pp. 1-10, 2020.

[23] H. Cunningham, "GATE, a general architecture for text engineering," *Computers and the Humanities*, vol. 36, no. 2, pp. 223-254, 2002, doi: https://doi.org/10.1023/A:1014348124664.

[24] D. Maynard, K. Bontcheva, V. Tablan, N. Aswani, I. Roberts, and et al, *Developing Language Processing Components with GATE* 8, ed.: University of Sheffield, UK, 2014, p. 581. [Online]. Available: http://gate.ac.uk/sale/tao/index.html.

[25] F. B. Oliveira, A. Haque, D. Mougouei, S. Evans, J. S. Sichman, and M. P. Singh, "Investigating the Emotional Response to COVID-19 News on Twitter: A Topic Modeling and Emotion Classification Approach," IEEE Access, vol. 10, pp. 16883-16897, 2022.

Hamed Vahdat-Nejad is currently an associate professor at the computer engineering department of the University of Birjand. He received his Ph.D. from the computer engineering department of the University of Isfahan in 2012, his M.Sc. degree from the Ferdowsi University of Mashhad in 2007, and his B.Sc. degree from the Sharif University of Technology in 2004. He was a research scholar with the Middleware laboratory at the Sapienza University of Rome in the 2011-2012 period. Currently, his research is focused on smart city & IoT, Crowd-sourcing, and text processing. He has (co)published more than 50 papers in conferences and journals, and leads the Pervasive & Cloud computing Lab at the University of Birjand. He has served as the chairman of the 1st & 2nd International Workshop on Context-aware Middleware for Ubiquitous Computing Environments, "3rd, 4th, and 5th International workshop on Pervasive and Context-aware middleware" as well as 1st conference on healthcare computing systems and technologies. He has served as a TPC member for many conferences. Currently, he is an associate editor for Elsevier computers and electrical engineering journal.

Faezeh Azizi received her B.Sc., in Software Engineering from Bojnurd University in 2017. She is a Researcher of Information Technology Engineering in the field of "Natural Language Processing". She also had a substantial role in research team works in the area of "Topic Modeling and Sentiment Analysis" at the University of Birjand. She has published more than 5 papers in conferences and journals. She is currently pursuing the M.Sc. degree in Information Technology at the University of Birjand, Birjand, Iran, and she is an active member of the Pervasive & Cloud computing Lab there.

Mahdi Hajiabadi is currently a M.Sc. student at the computer engineering department of the University of Birjand. He received a M.Sc. degree in computer software engineering from the University of Birjand in 2019. He is currently working in the field of Natural Language Processing.

Fatemeh Salmani is a graduate student of Information Technology Engineering at Birjand University. She received her B.Sc. degree in 2015 from Damghan University. he has served as a member of the Pervasive & Cloud computing Lab at Birjand University Since 2019. His research interests include Natural Language Processing, Machine Learning, Sentiment Analysis, and Text Mining.

Sajedeh Abbasi is currently an excellent postgraduate student at Birjand University. She received her M.Sc. degree from Birjand University in 2018. His current research focuses on smart cities, the Internet of Things, crowdsoursing, and text processing.

Mohadese Jamalian received the B.Sc. degree from the Birjand University of Technology in 2019. She is currently pursuing the M.Sc. degree in the University of Birjand. Her research interests include text processing and data mining.

Reyhane Mosafer received the B.Sc. degree from the Birjand University of Technology in 2019. She is currently pursuing the M.Sc. degree in the University of Birjand. Her research interests include text processing, image quality assessment and visual saliency detection.

Hamideh Hajiabadi is currently a researcher at Karlsruhe Institute of Technology, Germany. She has been working in machine learning area since 2014 and her main focus is now on deep networks and their theory and applications.

Wathiq Mansoor is a Professor at University of Dubai. He has an excellent academic leadership experience in well-known universities worldwide. He earned his Ph.D. in computer engineering from Aston University in UK. His doctoral work was on the design and implementations of multiprocessors systems and communications protocols for computer vision applications. He has published many research papers in the area of communication networks, Intelligent Systems, ubiquitous computing, web services, and neural networks. His current research is in the area of innovation in technology and management and smart cities focusing on Internet of Things and communications infrastructure. He has organized many international and national conferences and workshops. He is an executive member of IEEE UAE section.