A R T I C L E

Individual mobility pattern in Malaysia during COVID-19 Recovery Movement Control Order partial lockdown

Uznir Ujang | Suhaibah Azri

3D GIS Research Lab, Faculty of Built Environment and Surveying, Universiti Teknologi Malaysia, Johor Bahru, Malaysia

Correspondence
Uznir Ujang, 3D GIS Research Lab, Faculty of Built Environment and Surveying, Universiti Teknologi Malaysia, 81310 Johor Bahru, Johor, Malaysia.
Email: mduznir@utm.my

Abstract

Malaysia Recovery Movement Control Order (RMCO) aims to bring the business, education, tourism and other industry sectors back into operation. Due to movement constraints that result in local economic patterns, individual mobility patterns are expected to occur. However, this matter needs further investigation from people's spatial behaviour during the RMCO. Therefore, this research proposed a new technique for analysing people's spatial behaviour patterns via geo-tagged data. The data from social media users are gathered using data mining techniques. Geographical Information System (GIS) is used to show the geolocation of social media users and analyse their spatial behaviour. The finding of this analysis shows higher levels of movement recorded when the RMCO was enforced; a distinctive pattern where spatial trajectory length is high but spatial area coverage is low. It is noticed that the focal points are concentrated in urban areas and tourism attractions.

K E Y W O R D S
COVID-19, geo-tagged data mining, individual's mobility, Recovery Movement Control Order, spatial behaviour

1 | Background

A new virus outbreak was reported in Hubei province, Wuhan, in December 2019. According to Gorbalenya et al. (2020), this virus is a type of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and is identified as novel beta-coronavirus or 2019-nCoV (COVID-19). Although the world has encountered a few infectious diseases such as Ebola, influenza A (H1N1), SARS, MERS and Zika in the past two decades (Reperant & Osterhaus, 2017), the new COVID-19 virus appears to have had a significant impact on the world's economy, resources and human health (Chakraborty & Maity, 2020; He et al., 2020). Even though China had taken immediate action and measures to control the outbreak, the virus rapidly spread to other countries worldwide within a few months. On 11 March 2020, World Health Organization (WHO) announced the COVID-19 outbreak as a pandemic disease. As of 23 June 2020, 9,071,984 people worldwide had confirmed COVID-19 infection, and 470,716 deaths were recorded in 113 countries. According to John Hopkins (2020), the countries most affected were the United States, Russia, Brazil, Spain, Italy, France, Germany, the United Kingdom, India and China.

In response to the COVID-19 pandemic, Malaysia announced the 2020 Movement Control Order, known as MCO or PKP, on 18 March 2020 (Elengoe, 2020). The order was called a lockdown or partial lockdown, which barred Malaysians from overseas travel, restricted foreign visitors, shut down businesses and schools, introduced inter-state cross-border restrictions, and limited all activities, including religious activities, allowing only essential services (Sukumaran, 2020). The MCO was initially announced for two weeks, followed by an announcement of the first extension. Throughout the MCO, police conducted roadblock operations to monitor travellers and warn them to stay home and abide by the order.

However, the extent to which people adhered to this MCO was uncertain. Additionally, determining the individual mobility pattern within a country is impossible without large-scale data collection in every location. A novel strategy was required to address this issue, particularly during this pandemic. According to previous research, data acquisition on social media can represent the actual world situation. As such, we aimed to examine the movements of individuals during lockdown by examining spatial trajectories and locations of human attractions. Data mining techniques were applied to geotagged photos from social media platforms such as Instagram to accomplish the research objectives. The data collected were geocoded and georeferenced to be geolocated in the same way it would be in the actual world. Individual mobility patterns could be determined via this approach without the necessity for field data acquisition. This new approach could be used for future investigations to identify individual mobility patterns.

1.1 | Geotagged photos data mining

Digital photos nowadays are tagged with the location of the photo. These digital images are called geotagged photos. The location attached to a photo is the latitude and longitude coordinates of the Global Positioning System (GPS) coordinate system. Based on these coordinates, the location where the digital image was taken can be determined.

Numerous studies have used geotagged social media for a variety of goals. Geotagged activity, for example, can be used to track population mobility, refugee movement and disease dispersion. According to these studies, social media geotags have been shown to be a valuable tool for researchers interested in human activity and behaviour (Jurdak et al., 2015). Hawelka et al. (2014) have published global research results on mobility characteristics across nations, conducted using geotag data from Twitter. Twitter statistics are gathered to investigate the mobility profiles of different countries, including mobility frequency, radius, destinations and flow balance. Meanwhile, Cho et al. (2011) identified that human movement can be tracked by analysing mobility distance. Geotag data are analysed based on check-in locations on mobile phones. Additionally, social media serves as a platform for expressing the general public's opinions. Geo-tagged content on social media can also be used to make a person's presence known.

Comito (2021) recently published a study examining the role of social media, specifically Twitter, as an early indicator of the COVID-19 pandemic. The results indicate a significant correlation between tweets and COVID-19 data, indicating that Twitter may be a reliable indicator of virus transmission. Pandemic outbreaks may be monitored and comprehended better due to the data generated by user activity on social media platforms (Comito, 2021). This study evaluated how COVID-19 cases may be predicted using social media.

Nowadays, social media serves as a platform for expressing our views on global issues, political issues and personal experiences. Social media can monitor and assess current events and sentiments in this instance. For example, Alvino (2020) developed a tweet mining system based on geotagging tweets to monitor, analyse and visualise negative citizen sentiment. Besides that, geotagged data are also used for environmental monitoring, such as earthquake detection (Sakaki et al., 2010). They examine the location of tweets to determine the approximate location of the earthquake. Researchers also use geotagging for smart tourism. Nguyen et al. (2017) collected geotagged cultural heritage resources from social media platforms and facilitated visitor participation. Additionally, Zheng et al. (2012) conducted research in which they mined geotagged photographs from Flickr and Panoramio to gather and analyse tourist movement patterns depending on tourist attractions. This research demonstrates how the benefits of geotagged photos may be utilised to analyse tourist movement patterns and regulate traffic flow in congested tourist attraction locations. Okuyama and Yanai (2013) also used the same technique to get the best route suggestions to develop a travel route system. This system can also be used to avoid traffic congestion around central attractions. Further advanced research can also be found in studies by Ameen et al. (2020) and Zhong et al. (2020).

Cesario et al. (2017) stated that mobile devices are becoming more common, making it easier to collect data about how people move. These data can then be used to discover what people do. Among the things this paper covered was TPM, an algorithm that helps with the whole process of finding out how people move around. The algorithm is divided into two parts: first, it identifies dense zones and locations with a high density of individuals passing through; and second, it

extracts trajectory patterns from these dense areas. As demonstrated, the information gleaned from social media posts can be used to conduct additional scientific research.

Meanwhile, Comito et al. (2019) suggested a method for clustering topics from user-generated social media messages. The algorithm uses a novel measure of similarity that incorporates both syntactic and semantic information inherent in the text. The post's latent feature vector representations are created by training the Word2Vec skip-gram model on a vast corpus to improve the check in-topic mapping learned on a smaller corpus. The word vectors generated via neural embedding reveal linguistic regularities and patterns that enable significant accuracy increases over purely syntactic techniques and state-of-the-art technologies. This research suggests that future research should use larger datasets from various contexts.

Social media platforms make a range of data types available, one of which is geographical data. These data may be used for various applications, including mobile research, studying human behaviour, responding to natural disasters, and planning. The critical component is linking geographical data with geotagged data. According to Flatow et al. (2015) and Oleksiak (2014), not all data from social media are geotagged. Consider mining a large volume of data, but only a tiny percentage of it is geotagged. Manipulating these data might be difficult for data miners, mainly for trajectory reasons. Additionally, geotagged data taken from social media might be erroneous or misleading (Schlosser et al., 2021). For example, a study by Middleton et al. (2018) shows that geotag data are projected kilometres away when matching was done in OpenStreetMap. According to Schlosser et al. (2021), the missing location of geotagged data can be dealt with through three techniques: geocoding, geoparsing and geotagging. These techniques are based on data extraction. Consider the case where geographic information is not included in geotagging data. In such an instance, it may still be derived from the text's contextual location (e.g., Prime Minister was giving a speech in Washington). In this case, the location information contained in the phrase 'Washington' can still contribute value to the geotag data.

Additionally, geocoding can assist in locating a missing place by converting addresses and postal codes to their corresponding geographical coordinates (latitude and longitude). While geocoding accuracy is debatable, research by Faure et al. (2017) indicates that the accuracy is appropriate for various applications. Due to the advantages and demand for geocoding, commercial geocoding Application Programming Interface (API) services such as GeoIQ and Geocoding by ESRI are now available to meet customer demands. According to prior research and studies, geotag and geocoding methodologies are critical for assessing social media data. Thus, both geotagging and geocoding strategies are used in this work throughout the pre- and post-processing stages of data mining.

Based on previous research, it can be seen that geotagged photos uploaded on social media can be used to analyse individuals' mobility patterns and subsequently be able to see the rate of community compliance to government directives for the RMCO situation. Most data mining is conducted using social media platforms such as Twitter, Panoramio and Flickr. However, a survey conducted by the Global Web Index found that 70% of internet users in Malaysia use Instagram to interact and update their activity, compared with 40% who use Twitter. According to the Internet Users Survey 2020, Instagram users increased to 63.1%, while Twitter users increased to 37.1% in 2020 (Commission, 2021). However, no study has been conducted in Malaysia that utilises geolocation data from Instagram to determine an individual's mobility. Combining data mining techniques on the Instagram social media platform with geographical analysis in the geospatial field makes it possible to map an individual's movement and use it as a reference for those in authority. And this demonstrates that, in the future, social media data such as Instagram can be used for further scientific research.

Using these geotagged photos, this research uses data mining techniques to obtain user location information (from social media) during Malaysia's Recovery Movement Control Order (RMCO) partial lockdown. The advantages of using data mining techniques in social media can be seen in previous research conducted by Barbier and Liu (2011). However, the geotagged photos information used in this research was then used to see the movement of people after the MCO was withdrawn and replaced with an RMCO. This movement pattern is measured by looking at the spatial trajectories of people, and the spatial area covered throughout the implementation of the RMCO. It is believed to help those responsible agencies and parties in controlling and refining enforcement in curbing the COVID-19 pandemic from spreading further. The following section will describe the MCO and RMCO implemented in Malaysia and the details of movement control or movement limitations for people in Malaysia.

1.2 | Movement control order

The Prime Minister of Malaysia formally announced the MCO on 16 March 2020 (PMO, 2020). The Malaysian government decided to implement a nationwide MCO from 18 March 2020 to 31 March 2020. Under the Prevention and Control of Infectious Diseases Act 1988 (Act 342). This MCO aimed to reduce the spread of infection locally (Khor et al., 2020).

After two weeks of enforcement of the MCO, new cases continued to increase with occasional plateauing, hence the MCO was extended from 1 April to 28 April 2020. The RMCO was the sixth phase, effective from 10 June to 31 August 2020. Table 1 shows the phases involved during the implementation of the MCO.

This sixth phase (RMCO) was the recovery phase from the impacts of COVID-19 pandemic. The Malaysian government's initiative to address the COVID-19 pandemic was successful in phases 1–5 of the MCO. During these phases, the number of cases of COVID-19 infection in the country were successfully reduced from hundreds of new cases per day to one-digit cases per day. Among the rules and regulations that were gazetted during these five phases of the MCO were the following (Time, 2020):

1. Public gatherings and movements are entirely restricted, including religious, sports, social, and cultural activities. All places of worship and business premises must be closed to enforce this restriction, except for supermarkets, public markets, and convenience stores selling daily necessities.
2. Malaysians travelling abroad are subject to various restrictions. If they have returned recently from a foreign trip, they must undergo a 14-day health examination and voluntary quarantine.
3. All tourists and foreign visitors are subject to entry restrictions.
4. All kindergartens, public and private schools, and other educational institutions will be closed.
5. The closure of all public and private institutions of higher education.
6. Closure of all public and private establishments except those providing essential services to the country.

After almost three months of movement restrictions, nearly all sectors, including the economic and education sectors, were allowed to operate again under the Standards of Procedure (SOP) set by the National Security Council (MKN). After the travel restrictions were lifted, people were seen moving freely in the country and could carry out activities that were not previously permitted during the MCO. However, there were concerns about a second COVID-19 wave resulting from this movement. Consequently, monitoring by the authorities was necessary to see the people's self-control on the SOP that had been established. Enforcement agencies were required to identify the spatial trends of people's movements after a lockdown or partial lockdown has been withdrawn.

General observations indicated that most people took the opportunity to do inter-state movement activities for domestic tourism, social gatherings by visiting exciting places, and recreational activities such as cycling and hiking. As a result, this paper attempts to look at the spatial behaviour pattern of people's movements following the removal of these movement restrictions. This research aims to see which industries or sectors were the people's focus after months of not travelling beyond a 10 km radius of their respective homes. In this way, SOP compliance monitoring can be carried out in areas known to be hot spots for people in Malaysia to visit, and we believe this approach can be a reference to other countries that are still in the lockdown phase.

2 | METHODS

During MCO, social media has become the central platform to acquire information on the COVID-19 pandemic and keep people connected through social interaction, entertainment and shared information. Social media data mining was used to conduct this research. Social data mining is a relatively recent development in data mining. The data are

TABLE 1 MCO phases

MCO	Name	Duration
Phase 1	MCO	18 March–31 March 2020
Phase 2	MCO	1 April–14 April 2020
Phase 3	MCO	15 April–28 April 2020
Phase 4	MCO CMCO	29 April–3 May 2020 4 May–12 May 2020
Phase 5	CMCO	13 May–9 June 2020
Phase 6	RMCO	10 June–31 August 2020

Abbreviations: CMCO, Conditional Movement Control Order; MCO, Movement Control Order; RMCO, Recovery Movement Control Order.

filtered based on the geolocation information provided in social media posts. Indirectly, social media users' movements can be tracked and examined for experimental purposes. According to a survey conducted by the Global Web Index, 70% of internet users in Malaysia use the Instagram platform to communicate and update their activities, whereas just 40% of them use Twitter. Instagram is utilised as a platform in this research to acquire COVID-19 trajectory data in Malaysia. MCO and RMCO are two main topics that Instagram users are more likely to post during the pandemic on their social media accounts. It is possible to further examine this research by mining social media posts via often used hashtags, including the user's location.

The people's mobility patterns can be analysed using location-based data shared on social media. Social media users are now more likely to publish social media posts on present issues such as RMCO. These social media posts can be extracted along with the geolocation of the social media user by using commonly used hashtags. The social media user movement can then be carried out and analysed.

The overall diagram of the research workflow proposed in this study is illustrated in Figure 1. The first phase entails extracting data from social media platforms that users have shared. The retrieved social media data are based on the most widely used or popular hashtags in Malaysia related to RMCO for Malaysia. The hashtags used are #pkpp, #perintahkawalanpergerakanpemulihan, #rmco, #recoverymovementcontrolorder, #pkpp2020, #rmco2020, #normabaharu, #hapuscovid19, #newnorm, #newnorms and #normabaru. The data are extracted from public social media posts together with their metadata. As an overview, the total number of posts on 19 July 2020 is illustrated in Table 2.

Based on Figure 1, the following phase involves acquiring the location ID for social media data that have been successfully extracted. This technique is used to determine the geographical location of each shared post. Meanwhile, geocoding
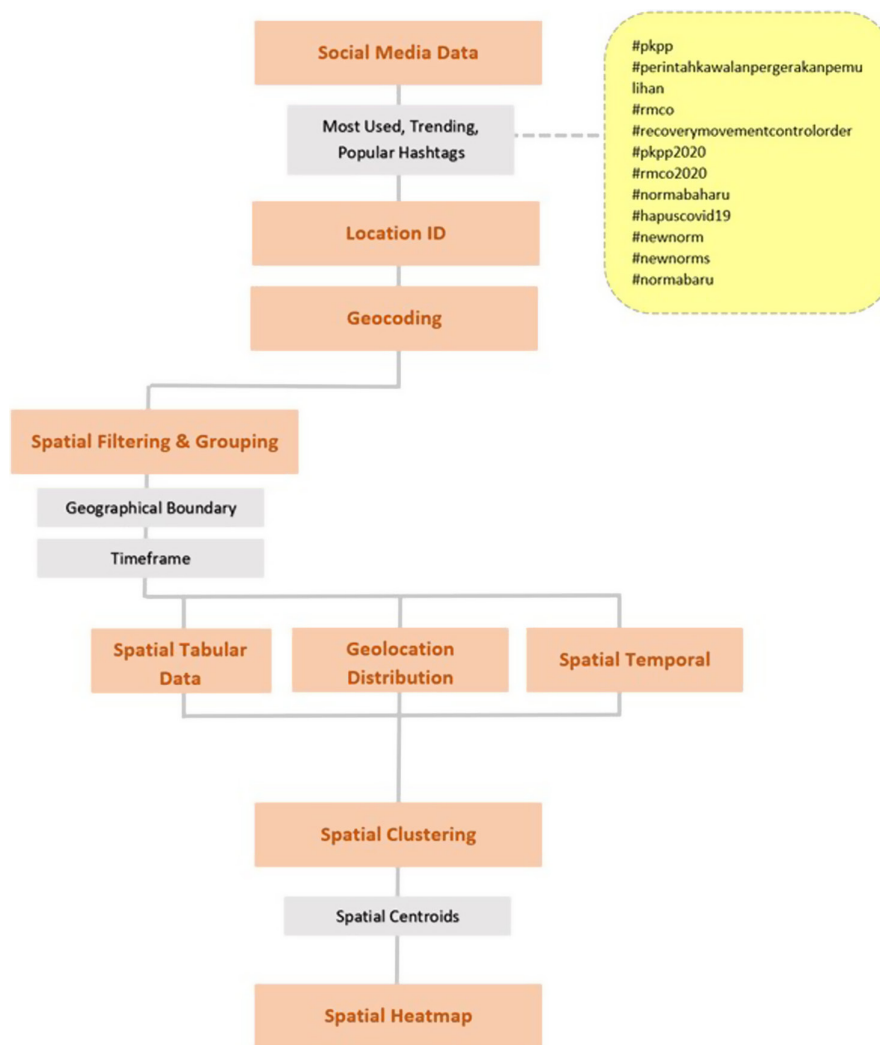

FIGURE 1 Overall research workflow in determining individual mobility patterns in Malaysia during COVID-19 Recovery Movement Control Order (RMCO)

TABLE 2 Total hashtags posts on 19 July 2020 at 18:08

No	Hashtag	No of posts
1	pkpp	150,011
2	perintahkawalanpergerakanpemulihan	6859
3	rmco	31,488
4	recoverymovementcontrolorder	1882
5	pkpp2020	5907
6	rmco2020	364
7	normabaharu	18,456
8	hapuscovid19	78,310
9	newnorm	68,468
10	newnorms	4727
11	normabaru	9225

Source: www.instagram.com

is the next phase, followed by the spatial filtering and grouping approach. Nonetheless, the number of posts extracted from the social media posts in this study is not equal to the number of posts presented in Table 1. This is because specific social media posts do not contain the geolocation of the post in its metadata; it depends on whether the social media user provides the current geographical location position. Therefore, only social media posts with their geographical location are used in this research.

Next, social media posts (along with their geolocation) go through the geocoding process. Geocoding is a process of putting the objects in the real-world coordinate system (in this situation, the geolocation of the social media posts). The coordinate system used in this research is the WGS84, a coordinate system used in GPS.

Once the geocoding process has been completed, the social media data are filtered based on geographical borders. The goal is to restrict the data extracted to the geographical boundary of Malaysia. Although the hashtags used are based on RMCO-related posts in Malaysia, several posts are located outside of Malaysia's geographical boundary, not associated and geo-positioned beyond the scope of this research. The timeframe for each posting, on the other hand, is used to group the data based on the RMCO phases.

Referring to Figure 1, once the data collection procedure has been completed (spatial filtering and grouping), information such as spatial tabular data, geolocation distributions, and spatial-temporal information can be generated. This information is critical since it is used as input for the spatial clustering process, generating spatial centroids prior to generating the spatial heatmap. A spatial heatmap is a data visualisation approach that uses colour to depict the magnitude of a phenomenon in two dimensions. The area in focus has a higher colour intensity than the area with low colour intensity. These data identify user emphasis areas based on social media post sharing.

3 | FINDINGS AND DISCUSSION

Statistics of social media posts extracted using hashtags are shown in Table 3. Overall, 69.40% of the social media posts' metadata were extracted. In other words, almost 70% of the total number of posts contain geolocation information in their metadata. The highest three refer to the hashtags #perintahkawalanpergerakanpemulihan (81.08%), #hapuscovid19 (79.10%) and #normabaharu (77.34%). In terms of the highest number of posts, the hashtags are #pkpp (105,901 posts), #hapuscovid19 (61,947 posts) and #newnorm (42,103 posts). However, the time spent extracting the social media posts was 95.14 h.

The extracted metadata were then filtered using the geographical boundaries parameter to make them valid for the situation in Malaysia. Although the hashtags used referred to the situation in Malaysia, the possibility existed that users outside Malaysia would use them. Therefore, without any filtering, further spatial classification would be affected by the inaccuracy of the data provided.

Next, the metadata were used in the process of geocoding, which is vital for locating and georeferencing the position for every social media post. Based on the number of georeferenced social media posts, 20.8% or 54,205 posts were

TABLE 3 Social media data statistics based on hashtags

No	Hashtag	No of posts	Metadata extracted	%
1	pkpp	150,011	105,901	70.60
2	perintahkawalanpergerakanpemulihan	6859	5561	81.08
3	rmco	31,488	18,104	57.49
4	recoverymovementcontrolorder	1882	1118	59.40
5	pkpp2020	5907	2947	49.89
6	rmco2020	364	142	39.01
7	normabaharu	18,456	14,273	77.34
8	hapuscovid19	78,310	61,947	79.10
9	newnorm	68,468	42,103	61.49
10	newnorms	4727	2841	60.10
11	normabaru	9225	5788	62.74
	Total	375,697	260,725	69.40

TABLE 4 Social media posts extracted based on geo-location ID

No	Hashtag	Geo-boundary
1	#pkpp	26,329
2	#perintahkawalanpergerakanpemulihan	590
3	#mco	5582
4	#recoverymovementcontrolorder	512
5	#pkpp2020	1420
6	#rmco2020	58
7	#normabaharu	2586
8	#hapuscovid19	10,819
9	#newnorm	4368
10	#newnorms	265
11	#normabaharu	1676
	Total	54,205

recognised as bounded in Malaysia's geographical boundary. Table 4 presents further details of the number of social media posts based on the hashtags used.

Moreover, the geolocation distribution of social media users can be produced based on the geocoding procedure that was previously done during the data processing phase. Figure 2 depicts the geolocation distribution for social media users in Malaysia after lockdown or during the RMCO phase. This information was retrieved after all of the data had been mined, geocoded and filtered to determine the geolocation of social media users within the geographical boundaries of Malaysian territory. The geolocation distribution is concentrated on the west coast of peninsular Malaysia and major cities on the east coast of peninsular Malaysia and Borneo Island (Sabah and Sarawak).

Based on the data obtained, it was found that 11,774 users of social media accounts actively stated their geolocation. However, out of that number, not all social media accounts regularly post their status on social media. Based on Figure 3, User ID8402532061, ID36570252971 and ID15423612909 occupy the top social media status with 555, 320 and 281 posts, respectively. The data organisation in Figure 3 is by the name (ID) column and the total number of postings (Length) column. The following diagram (Figure 4) shows the record of social media status daily throughout this RMCO. The data used in this research cover the time between 9 June 2020, 00:01:11 and 20 July 2020, 04:37:45.

Another finding concerns the geographical extent of social media users. This procedure maps the geolocation of social media users daily. Figure 5 shows the geolocation distribution of users based on the WGS84 coordinate system published
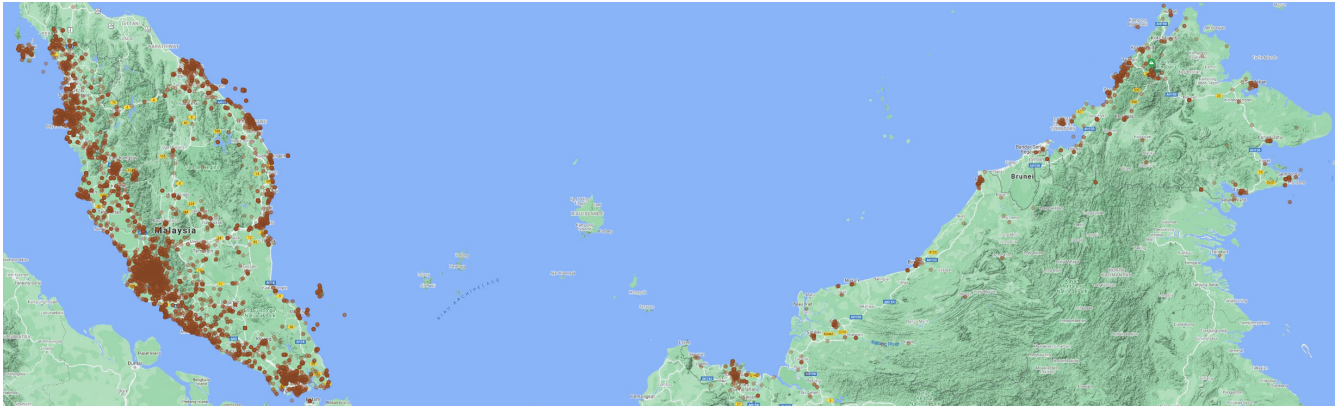
FIGURE 2 Geolocation distribution of social media users during Recovery Movement Control Order (RMCO) partial lockdown

FIGURE 3 Social media user IDs and number of posts with data length and data type information. RMCO, Recovery Movement Control Order
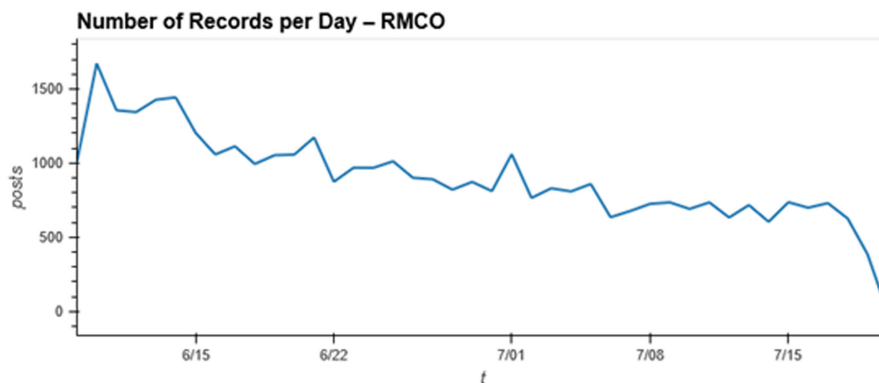

FIGURE 4 Social media records obtained between 9 June 2020, 00:01:11 and 20 July 2020, 04:37:45. RMCO, Recovery Movement Control Order
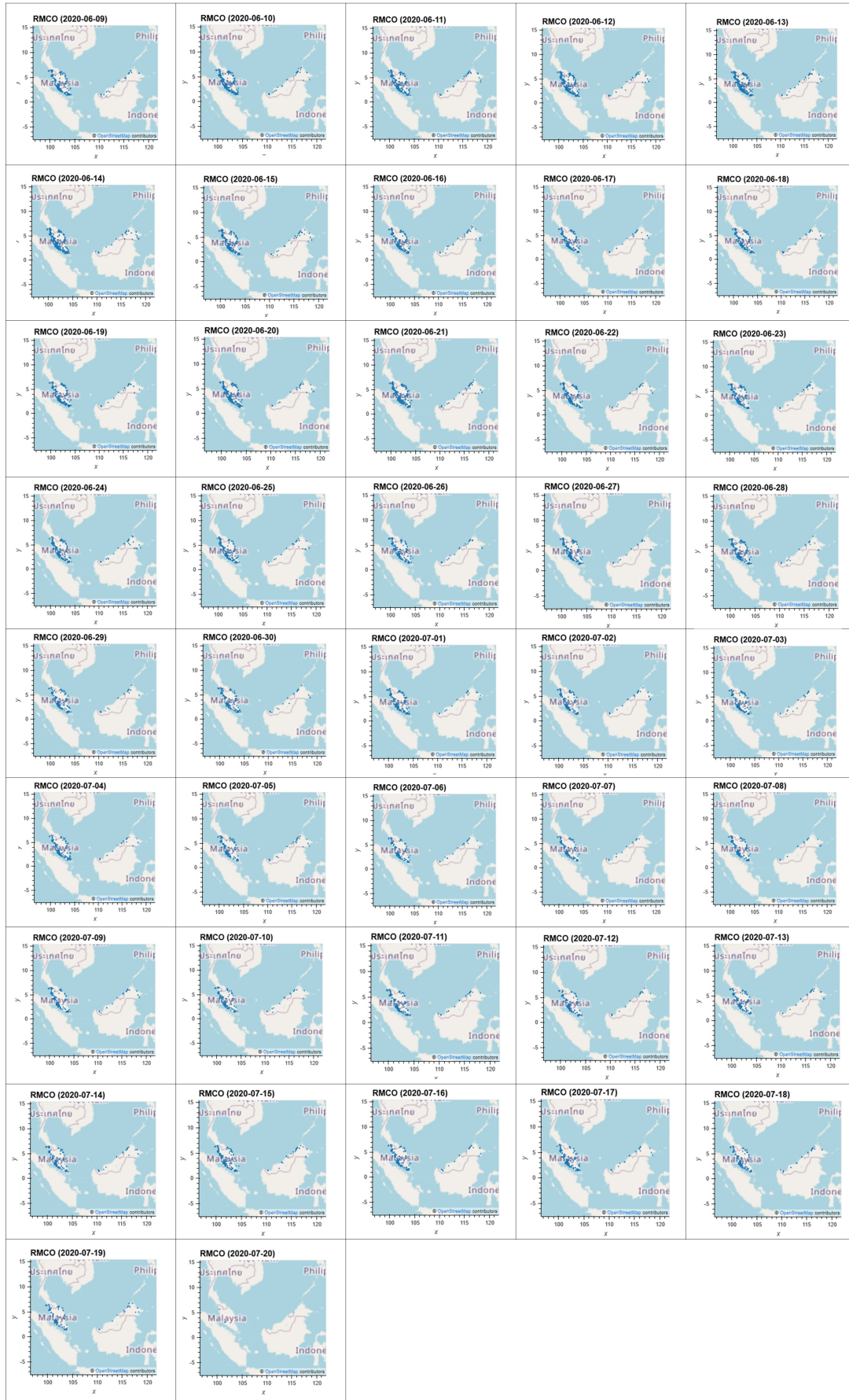
FIGURE 5 Geographical extent of social media users during Recovery Movement Control Order (RMCO) partial lockdown

during the previous geo-reference process. With this mapping and a map scale of 1:65,000,000, the location of social media users can be posted on the map, and the focus of the location (hot spots) of the social media users can be seen visually. This simplifies the process of seeing the geolocation trends of social media users (spatial-visual).
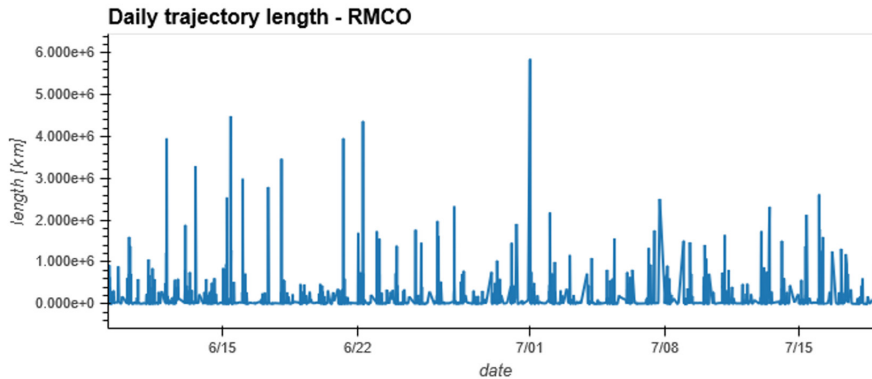
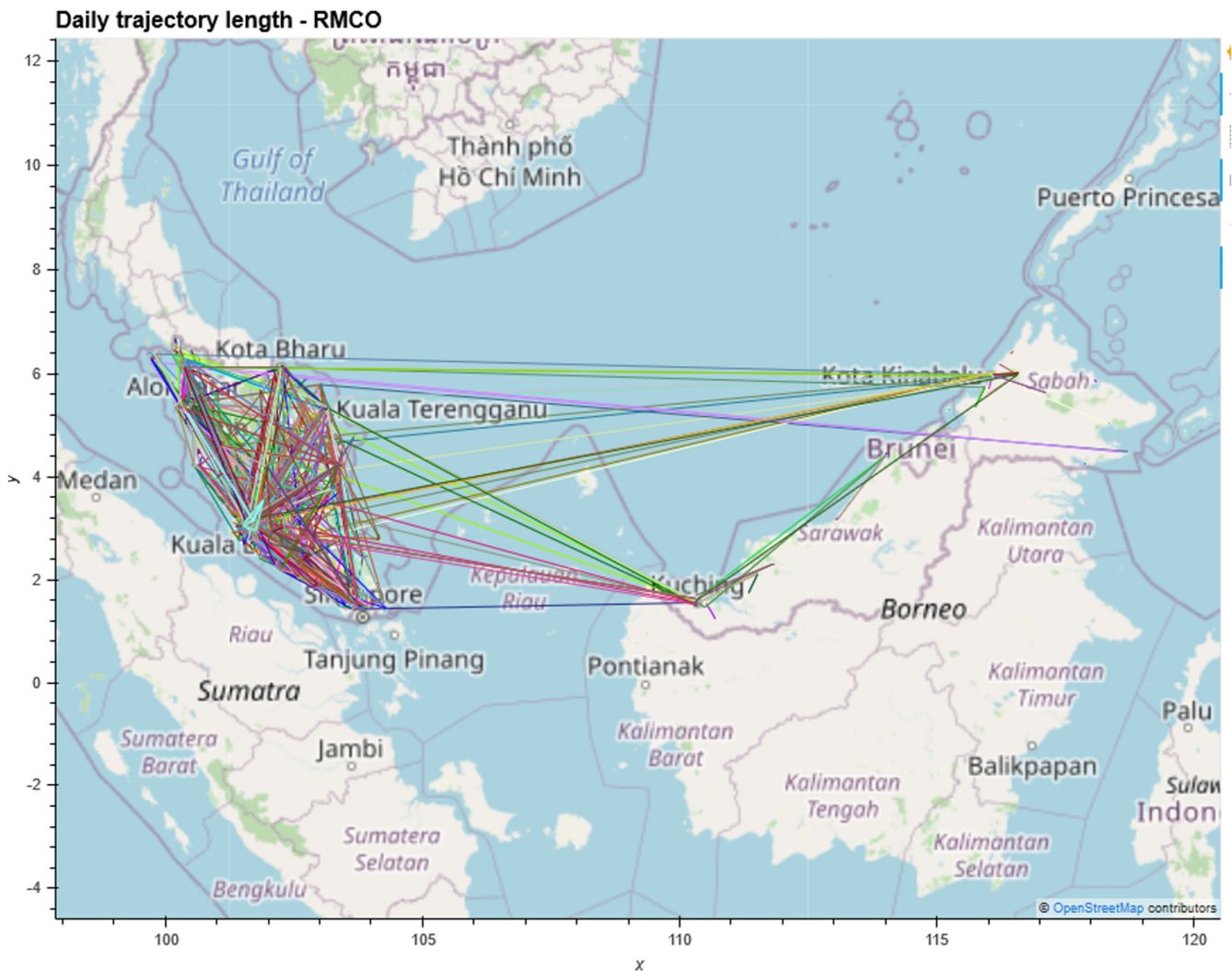**FIGURE 6** Spatial trajectory length. RMCO, Recovery Movement Control Order

FIGURE 7 Spatial trajectory map. RMCO, Recovery Movement Control Order

Based on the illustration shown in Figure 5, it can be seen that the geolocation trend of users is on the west coast of peninsular Malaysia. At the same time, there is also a visible concentration on the east coast of peninsular Malaysia, especially in major cities of the states of Kelantan, Terengganu and Pahang. In comparison, fewer hot spots can be seen in the Borneo archipelago, Sabah and Sarawak. Furthermore, this distribution appears to be consistent throughout the RMCO. However, on the last day of RMCO (20 July 2020), a significant decrease can be seen in mapping the geolocations of social media users. The data sources obtained during the data mining process on that specific date are limited. The lack of data resulted in geolocation mapping for social media users on that day that did not reflect the actual situation of the RMCO.

Next, the mined social user data were processed in spatial trajectories. By definition, the spatial trajectory is the path of an object with mass in motion that follows through space as a function of time. The trajectories for each User ID were examined and constructed based on the records of processed social media posts. However, some users publish posts but do not have any trajectory. This is possibly due to the lack of trajectory/path for the position of the posts published, or the number of posts is not adequate for the user (i.e., one social media post).

A total of 2289 spatial trajectories were created. Figure 6 shows the length of daily spatial trajectories for the social media users of this study. Figure 7 is a spatial trajectories map that visually displays the movement of social media users. This movement is mapped using Euclidean distance for each spatial trajectory. Social media users are displayed in different colours to show data already being mined.

Based on the spatial trajectories shown in Figures 6 and 7, it was found that the movement of social media users increased when the RMCO was first enforced. This is probably due to the restricted user movement limit when the MCO (before RMCO) was retracted. Therefore, social media users show an increasing value for their respective spatial trajectory lengths. An increase can be seen from the first week of RMCO implementation until the end of the second week. A sharp increase occurred at the end of the third week, which doubled compared with previous weeks. Interestingly, a drastic decline started on the fourth week and beyond. Daily spatial trajectories for social media users are smaller in weeks 4, 5 and 6, averaging around less than 1^{e+6} km length.

Yet despite the results obtained for daily spatial trajectory lengths, differences occurred during subsequent analysis. This research further investigated the area covered (km^2) by social media users during RMCO. Figure 8 shows the daily coverage area of spatial trajectory patterns by social media users during RMCO.

Based on Figure 8, it was found that there was an increase at the end of the first week of RMCO. The highest values were found approaching 50 km^2. A similar pattern was found for the second week, where the spatial value of the covered area increases at the end of the week. Looking back at the spatial trajectory lengths in Figure 6, the third week has the highest recorded values. However, the third week showed a low spatial value for the covered area. It clearly shows social media users' high movement (length) but low spatial covered area. The social media user movement does not seem to be widespread. Subsequently, the fourth week and onwards showed an insignificant increase in the daily covered area obtained. Figure 9 illustrates a heatmap of people's spatial behaviour during RMCO. As can be observed from the heatmap, most of the population appeared to be concentrated in major cities and locations with a high level of tourism activity.

Based on the findings of this study, it can be proved that the technique proposed in this study can be used as one of the methods for studying individual mobility patterns during movement restrictions, either during a pandemic or any related regulations. The data mining method shown in this study can be applied, and then the geographical location of the data can be identified. Based on this approach, spatial grouping and filtering can generate spatial tabular data, geolocation
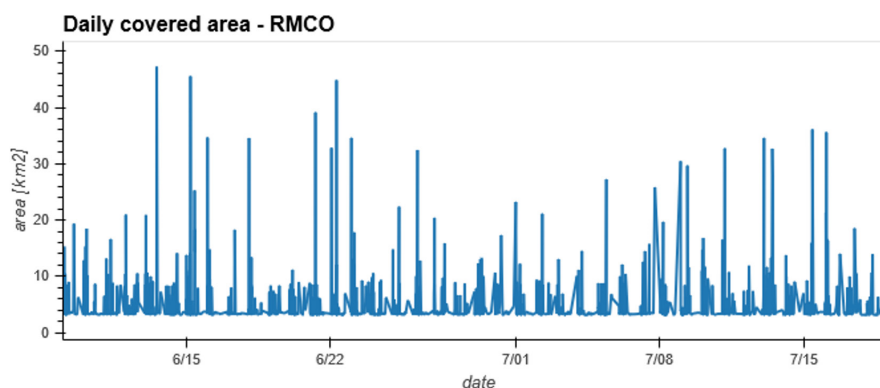

FIGURE 8 Spatial covered area. RMCO, Recovery Movement Control Order
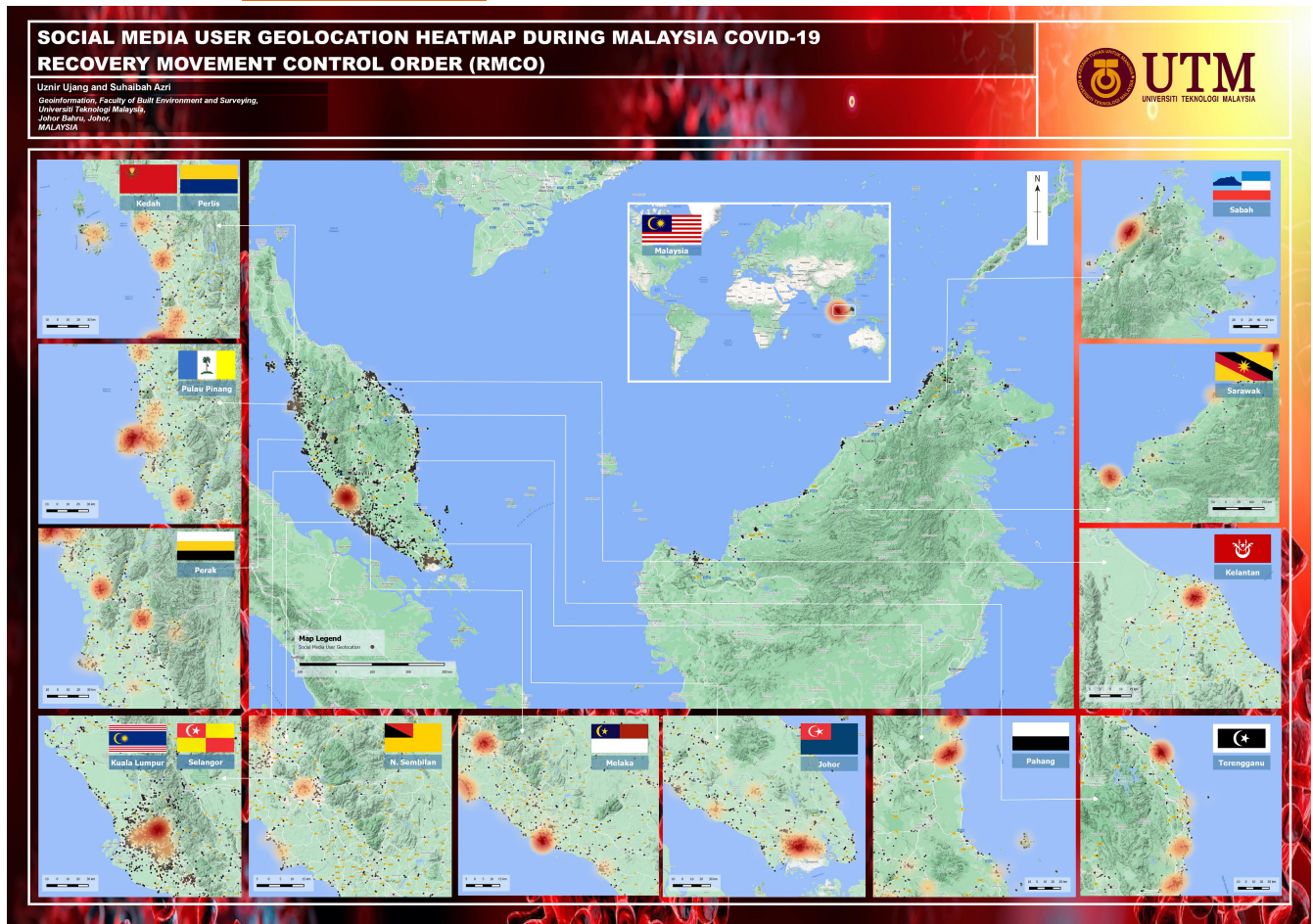
FIGURE 9 Heatmap showing people's distribution during Recovery Movement Control Order (RMCO) partial lockdown

distributions, and spatial-temporal information. This information would be beneficial in spatial analysis involving movement (temporal). Based on this research, a detailed spatial heatmap was constructed for each state in Malaysia.

4 | CONCLUSION

The RMCO enforced by the Malaysian government aimed to curb the spread of the COVID-19 pandemic in the country and, at the same time, resume the country's economic growth. However, the war against the COVID-19 pandemic is not over. Therefore, the flexibility provided during the RMCO still needs to be controlled and protected. It avoids the risk of transmission and the increase in COVID-19 patient numbers. Therefore, to some extent, this study shows people's spatial behaviour after the MCO was eradicated. Work and social-related activities, including religious activities, were allowed to resume according to the standard of procedure (SOP) determined by the National Security Council (MKN). This research used the latest data mining techniques for social media users. Geotagged data of social media users were processed and analysed to identify their spatial behaviour. GIS technology allowed these movements to be identified and analysed by looking at user focus areas, spatial trajectories and spatial areas covered. Daily geotagged mapping of data showed that social media users were concentrated in the areas around the west coast of peninsular Malaysia, in major cities as well as tourist hotspots. Initial findings found that an individual's mobility might occur towards these specified areas. This information is helpful for the authorities and can help control the spread of the COVID-19 pandemic. Although a decline in the number of COVID-19 positive patients has been shown, the threat of this pandemic still exists and we need to be on alert. Additionally, our study findings can be used by other countries when dealing with the pandemic with the enforcement of an MCO or an approach that is similar to Malaysia. It is believed that enforcement and monitoring can be done systematically and efficiently by identifying and understanding people's behaviour.

4.1 | Limitation

The limitations of conducting this research are mentioned herein. The geolocation accuracy of social media users in this study depended on the accuracy of the positioning or GPS found on the user's mobile phone. The difference in the precision of the GPS device used may give some differences in the accuracy of the social media user's position. In addition, the trajectory length generated is based on measurements derived from the Euclidean distance. There will likely be differences in distances or areas that social media users have been through in the analysis of this study. On the other hand, the username of the social media account is anonymous to protect the privacy rights of every user of the social media platform. This is because this research has no intention of tracking social media users, but to analyse the user movement pattern during the RMCO phases.

CONFLICT OF INTEREST
The authors declare that they have no competing interests.

DATA AVAILABILITY STATEMENT
The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID
Uznir Ujang https://orcid.org/0000-0001-5281-8478
Suhaibah Azri https://orcid.org/0000-0001-7926-9608

REFERENCES
Alvino, N. (2020). Sentiment Analysis and VisualFation for Citizen Opinion in Social Media. Proceedings of The National Conference on Undergraduate Research (NCUR) 2020, Montana State University.

Ameen, T., Chen, L., Xu, Z., Lyu, D., & Shi, H. (2020). A convolutional neural network and matrix factorization-based travel location recommendation method using community-contributed geotagged photos. *ISPRS International Journal of Geo-Information*, 9(8), 464. https://doi.org/10.3390/ijgi9080464

Barbier, G., & Liu, H. (2011). Data mining in social media. In C.C. Aggarwal (Ed.) *Social network data analytics*. Boston, MA: US Springer, pp. 327–352. Available from: https://doi.org/10.1007/978-1-4419-8462-3_12

Cesario, E., Comito, C., & Talia, D. (2017). An approach for the discovery and validation of urban mobility patterns. *Pervasive and Mobile Computing*, 42, 77–92. Available from: https://doi.org/10.1016/j.pmcj.2017.09.006

Chakraborty, I., & Maity, P. (2020). COVID-19 outbreak: Migration, effects on society, global environment and prevention. *Science of the Total Environment*, 728, 138882. Available from: https://doi.org/10.1016/j.scitotenv.2020.138882

Cho, E., Myers, S.A., & Leskovec, J. (2011). Friendship and mobility: User movement in location-based social networks. Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining, San Diego, California, USA. Available from: https://doi.org/10.1145/2020408.2020579

Comito, C. (2021). How COVID-19 information spread in US: The role of Twitter as early indicator of epidemics. *IEEE Transactions on Services Computing*, 1. Available from: https://doi.org/10.1109/TSC.2021.3091281

Comito, C., Forestiero, A., & Pizzuti, C. (2019). Word embedding based clustering to detect topics in social media. 2019 IEEE/WIC/ACM International Conference on Web Intelligence (WI).

Commission, M.C., & a. M. (2021) *Internet users survey 2020*. Commission, M. C. a. M. Available from: https://www.mcmc.gov.my/skmmgovmy/media/General/pdf/IUS-2020-Report.pdf [Accessed 10th May 2021].

Elengoe, A. (2020). COVID-19 outbreak in Malaysia. *Osong Public Health and Research Perspectives*, 11(3), 93–100. Available from: https://doi.org/10.24171/j.phrp.2020.11.3.08

Faure, E., Danjou, A.M.N., Clavel-Chapelon, F., Boutron-Ruault, M.-C., Dossus, L., & Fervers, B. (2017). Accuracy of two geocoding methods for geographic information system-based exposure assessment in epidemiological studies. *Environmental Health*, 16(1), 15. Available from: https://doi.org/10.1186/s12940-017-0217-5

Flatow, D., Naaman, M., Xie, K.E., Volkovich, Y., & Kanza, Y. (2015). On the accuracy of hyper-local geotagging of social media content. Proceedings of the Eighth ACM International Conference on Web Search and Data Mining, Shanghai, China. Available from: https://doi.org/10.1145/2684822.2685296

Gorbalenya, A.E., Baker, S.C., Baric, R., Groot, R.J.d., Drosten, C., Gulyaeva, A.A., Haagmans, B.L., Lauber, C., Leontovich, A.M., & Neuman, B.W. (2020). Severe acute respiratory syndrome-related coronavirus: The species and its viruses – a statement of the coronavirus study group. 1-15. Available from: https://doi.org/10.1038/s41564-020-0695-z

Hawelka, B., Sitko, I., Beinat, E., Sobolevsky, S., Kazakopoulos, P., & Ratti, C. (2014). Geo-located Twitter as proxy for global mobility patterns [Article]. *Cartography and Geographic Information Science*, 41(3), 260–271. Available from: https://doi.org/10.1080/15230406.2014.890072

He, Q., Liu, J., Wang, S., & Yu, J. (2020). The impact of COVID-19 on stock markets. *Economic and Political Studies*, 8(3), 275–288. Available from: https://doi.org/10.1080/20954816.2020.1757570

Johns Hopkins University. (2020). COVID-19 Dashboard by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University (JHU) John Hopkin University. Available from: https://www.arcgis.com/apps/opsdashboard/index.html#/bda7594740fd40299423467b48e9ecf6

Jurdak, R., Zhao, K., Liu, J., AbouJaoude, M., Cameron, M., & Newth, D. (2015). Understanding human mobility from Twitter. *PLoS One*, 10(7), e0131469. Available from: https://doi.org/10.1371/journal.pone.0131469

Khor, V., Arunasalam, A., Azli, S., Khairul-Asri, M.G., & Fahmy, O. (2020). Experience from Malaysia during the COVID-19 movement control order. *Urology*, 141, 179–180. Available from: https://doi.org/10.1016/j.urology.2020.04.070

Middleton, S.E., Kordopatis-Zilos, G., Papadopoulos, S., & Kompatsiaris, Y. (2018). Location extraction from social media: Geoparsing, location disambiguation, and geotagging. *ACM Transactions on Information Systems*, 36(4), 1–27. Available from: https://doi.org/10.1145/3202662

Nguyen, T.T., Hwang, D., & Jung, J.J. (2017). Using geotagged resources on social media for cultural tourism: A case study on cultural heritage tourism. Cham, Switzerland: Lecture Notes of the Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering, LNICST.

Okuyama, K., & Yanai, K. (2013). *A travel planning system based on travel trajectories extracted from a large number of geotagged photos on the web*. New York, NY: The Era of Interactive Media.

Oleksiak, P. (2014). Analysing and processing of geotagged social media. *Information Systems in Management*, 3(4), 250–260.

Reperant, L.A., & Osterhaus, A.D. (2017). AIDS, Avian flu, SARS, MERS, Ebola, Zika... What next? *Vaccine*, 35(35), 4470–4474. Available from: https://doi.org/10.1016/j.vaccine.2017.04.082

Sakaki, T., Okazaki, M., & Matsuo, Y. (2010). Earthquake shakes Twitter users: Real-time event detection by social sensors. Proceedings of the 19th international conference on World wide web, Raleigh, North Carolina, USA. Available from: https://doi.org/10.1145/1772690.1772777

Schlosser, S., Toninelli, D., & Cameletti, M. (2021). Comparing methods to collect and geolocate tweets in Great Britain. *Journal of Open Innovation: Technology, Market, and Complexity*, 7(1), 44. Available from: https://doi.org/10.3390/joitmc7010044

Sukumaran, T. (2020). Coronavirus: Malaysia in partial lockdown from March 18 to limit outbreak. Available from: https://www.scmp.com/week-asia/health-environment/article/3075456/coronavirus-malaysias-prime-minister-muhyiddin-yassin

Time, N.S. (2020). Covid-19: Movement control order imposed with only essential sectors operating. Available from: https://www.nst.com.my/news/nation/2020/03/575177/covid-19-movement-control-order-imposed-only-essential-sectors-operating [Accessed 16th March 2020].

Zheng, Y.-T., Zha, Z.-J., & Chua, T.-S. (2012). Mining travel patterns from geotagged photos. *ACM Transactions on Intelligent Systems and Technology*, 3(3), 1–18. Available from: https://doi.org/10.1145/2168752.2168770

Zhong, L., Yang, L., Rong, J., & Kong, H. (2020). A big data framework to identify tourist interests based on geotagged travel photos. *IEEE Access*, 8, 85294–85308. Available from: https://doi.org/10.1109/ACCESS.2020.2990949