Personalized Route Recommendation Method for Field Survey Officers using Social Media Information and Administrative Border Maps

E Hardiyanto

Statistics Jawa Timur, Kendang Sari Street, 43-44, Surabaya, Jawa Timur

*Corresponding author's e-mail: eko.hardi@bps.go.id

Abstract. Changes in business processes in pandemic conditions are a must. The field survey were most affected, not only the interview process but also the route selection to survey location. To support the field survey officer, it is necessary to provide alternative route choices to the survey location as fast as possible. This research proposed a methodology that combines three information source, administrative border maps, google maps services, and information from social media that elaborated to provide best recommendation route to the assigned survey location. The combination of three different sources can enhance the current existing route that only relies on google map services. Our mechanism was tested on custom my maps application provided by Google and evaluated using system usability scale. This research aims to give the personalized route to field survey officers based on the assigned survey location and information from social media. The limitation of this research is that the social media channels used are still few, in the future, this research can be leveraged by integrating other platforms owned by the government and other public services to enrich the information.

1. Introduction

The current pandemic condition makes all activities carried out increasingly limited. The restrictions imposed by the Indonesian government called restrictions on community activities on a micro-scale (PPKM) make all sectors, both government, private, and community, have to change their way of doing business that is more adaptive to conditions. These restrictions make some areas inaccessible, from its restriction the activities that were affected are survey activities that require face-to-face interview contact with the respondents in the community.

Census and survey activities for basic data needs are carried out by the Statistics of Indonesia (BPS). Before the pandemic period the survey activity been done by BPS must conduct interviews in the field, while in the pandemic conditions the field activities had changed significantly to comply with the health protocol rules issued by the government. With restrictions on both micro and macro scales, survey activities must be adjusted immediately not only the business processes of the survey but also the field strategy. The problems that are currently being faced are the information on restriction and information on isolation from one location cannot be easily accessed, so it could be a factor that hinders survey officers from going to the location to carry out survey activities as usual. Others factor are the google route sometimes cannot reach the destination location of survey location because of lack of signal or other factors.

The implementation of this proposed route recommendation can increased success in the field of survey. The goals of the proposed mechanism with combination and integration between three different

sources such as administrative border maps from government, information from private companies such as Google has open access, then involving information from social media from community services will make the information more useful and accurate. The BPS activities almost using assigned survey locations in a specific area for each survey officer, using this research methodology by combining administrative border maps can give information for survey officers to minimize wrong scope area while collecting data. This research aims to give the personalized route recommendation to the survey officer to go to the assigned survey location.

2. Related Works

Many studies on big data involving social media as a source of information to identify disasters [1] location of ambulance services [2] inform about road congestion conditions [3] [4] [5] [6]. Social media also provides public APIs such as Twitter which are used to analyze transportation conditions in an area [4]. Social media lately can be used to do social information work such as helping to find missing children, helping to find blood donors, finding financial donors. In terms of personalized route, had been researched using big data from hot spot urban crime from government agencies [7]. By looking at the potential that exists, the information provided on social media is used to help determine paths to help solve problems in census or survey activities also give the personalized path to the survey officer.

Many activities for calculating routes and distances have been carried out, using the Euclidean distance, the shortest distance, using other methods that are tailored to the needs of the user and the most commonly used is google maps. Google maps itself are used as a tool in determining the accuracy of the location of the occurrence of problems or events [1] [8]. Google maps have been widely used in determining distances or spatial-based analysis [9] [10] [2]. Google map is a resource for maps that can be used for free by users. One of the features provided to users is that users can create personal and custom maps according to their individual needs [11]. Google map itself have often been used in providing road or route recommendations for users [12] [13], used in planning the location of electric charging stations for electric-powered vehicles [14] determining agricultural locations with satellite imagery [15], minimizing accidents [14] [3], to classify children's health [16] and determine the location of water pipe damage [8]. However, the use of google maps for officers' route recommendations were very limited and in this research, we proposed a new mechanism that can be proposed for the survey officer to minimize error scope and other problems in field survey caused by a restriction in location or because PPKM.

3. Methodology

3.1. Personalized Route Recommendation

This study uses data from the smallest administrative borders (SLS) map owned by BPS in conducting simulations. The administrative border data from BPS and from open administrative maps [17] were used to identify the start and end locations of survey officers. Furthermore, the data assignment of officers is simulated start and end with route maps obtained from google [18], then the information is combined with the results of extra information from social media Twitter [19]. The mechanism to select the finest route recommendation is showen in Figure 1.

In this study, a simulation will be presented using a search keyword from social media Twitter then the information that had been crawled will be extracted. The crawling and extraction process has been done using R software. The result of the information extraction process is used to provide the latest information regarding information traffic, PPKM, or other related keywords. Mapping information from social media to spatial form we use QGIS application software then converted to kml extension. To integrate all sources we use ‘my maps’ application provided by Google that is available for android and web access. Algorithm 1 showed detailed process and steps in this methodology. The algorithm gives detailed steps that survey officers to select a path that are more efficient they can pass by as show in the line 6 to 10 in the algorithm 1.

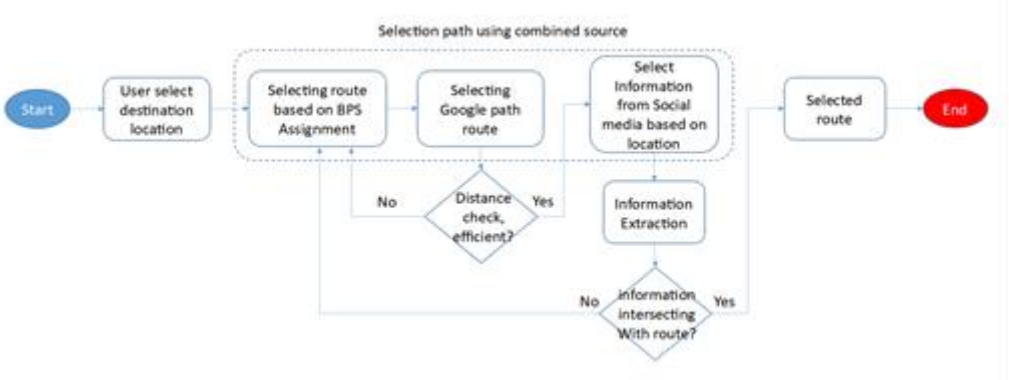

Figure 1. The route recommendation selection steps for field survey officer

This study proposes a methodology that combines data available in the government, namely data related to regional boundaries or administrative borders, and then integrated with google maps and social media information that has been extracted. Although research related to route selection recommendations is still very little, supporting research such as the use of google maps, then research on the use of social media has been widely carried out. So that with this research can be the basis for development to researchers.

Personalized Route Selection Based On Administrative Maps And Social Media Information
1. Start 2. Initialize location: location point of the user of the survey officer. • Assignment database for each survey location 3. Request: Assignment information data and location information are stored on the server. 4. Check administrative border maps: the location of origin and location of destination from the map data owned by BPS itself for the distribution of the initial location point and the point of a destination location for the survey officer. 5. Generate path: Path will be generated. 6. Select: Selection of the initial path is to use the closest available path from the location (All available paths) 7. Google maps route available: Then the next stage is determining the location of the shortest path that may be obtained by comparison users with the google path. 8. Social media information: Crawling information from social media 9. Information extraction: Select information related to the path or location based on related information such as traffic jams, restriction, or other things resulted from social media. 10. The result: social media information • If there's related social media information: Convert and match to layer maps social media information • Else: No information generated from social media 11. Display: all available information from administrative borders, google maps services, and social media information 12. End

Algorithm 1. Algorithm Personalized Route Selection based on administrative borders maps, google service and social media information

The limitation in this research is the addition of new limited sources of social media that are used only which have an open API such as Twitter. Information related to this location is based on the paths traversed by the paths that have been passed by the officers. Given the limitations in this test, we tried to combine the open-source API from Twitter and then with google path with kml format to determine

the length of the route as well as by calculating the nearest point on the map to determine the effective route from the origin to the destination point. Preparation of path selection by location to help speed up resolving the problem of delays in officers going to the location or officers will report the results of their work because there are obstacles during field implementation.

3.2. Comparison of existing google maps

The google maps services were map service that had widely used as the personalized route selection for almost all android users. In this research, we proposed to give more information for survey officer to select an alternative route to the assigned suvey location. The comparison of existing maps services and the proposed method that combine the information from administrative borders maps, the social media information and the route integration from google maps services were shown in table 1.

Table 1. Comparison existing google maps route and proposed method

Method	Advantages	Disadvantages
Existing Google maps	a. Easy to use b. User gives information traffic jam from other users that use and activate maps	a. Maps for assigned locations may be misleading if user didn't know the location b. Information about route only displayed on a route that passed by not all information were showed
Prosposed method (administrative border maps, google services and social media information)	Information on the user maps more detail: a. Location start and assigned survey location b. Information from social media (traffic jam, accident, PPKM and other) available	a. Require more time to provide information b. No specific application that integrate the information

4. Result and Discussion

4.1. Results

The testing of the proposed methodology has been carried out as shown in the Figure 2. The results of the simulation carried out using the proposed methodology from Figure 1. Figure 2 describes the steps taken in combining giving recommendations to a survey officer and show the testing had been done in this research. Not only the steps that have been done in the testing our methodology but also Figure 2 gives information the source that we use to get the source and explain the tools that used to convert or extract the information from social media crawling results to spatial format.

In the initial stage, the survey officers will be given an assignment by BPS in carrying out their activities. At the same time, information about starting point to destination location were crawled from social media as represented (number 1 in Figure 2). All information from social media were collected and filtered only related information that will be selected, we use R software to do the crawling process and information gathering related to the starting location, destination location, and the surrounding information (number 2 in Figure 2). The information from social media then been filtered and only select the most related information to support the selection route recommendation. The selected information then stored and extracted to get information such as PPKM location, street name, disaster location, location of an accident or other information that supported the survey officer to go the destination survey location safely (number 3 in Figure 2). The selected information then converted to kml format, the format that used by google platform.

The source comes from social media information generated from crawling data, the administrative border maps, and the google service maps that are combined using google custom maps service [18].

To import all the sources we converted to kml formatting maps, then imported in the website my maps google (number 4 in Figure 2). After all source maps have been imported using 'my maps' website, the user can download the android application to get the imported maps in mobile device format. The android application of 'my maps' only require small space in the phone so it can easily download by user. Using the android application, the imported maps layer can be view and the application can locate the current position of the user (number 5 in Figure 2).

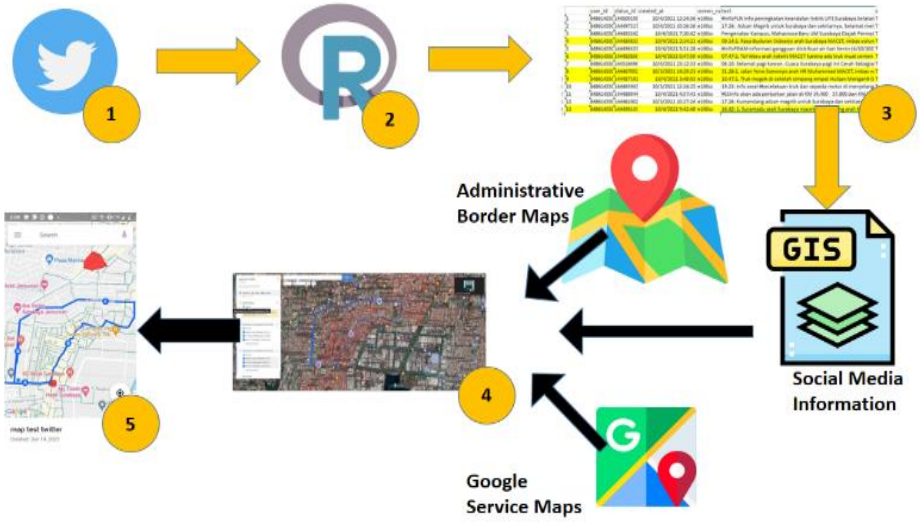

Figure 2. Personalized route recommendation system testing steps

The next step is to test if the maps provided using our proposed mechanism were useful. We tested using data of Surabaya, the Tenggilis Mejoyo sub districts. We tested to show starting point and endpoint that have been located in the same sub districs tenggilis mejoyo as shown in Figure 3. We elaborated using the information administrative border maps from statistics Jawa Timur and social media information that inform the current condition of street that around in the Tenggilis Mejoyo sub districs as shown in Figure 4. In the process to produce the information from social media we get the result from Twitter crawling text results using keywords that support the recommendation selection route for field survey officer. Some result from crawling process in Table 2 that indicate some of the resulted text that been crawled using R software. From Table 2 the next process is converting the information that relate to help the field survey officer converted to kml format using QGIS application. The process depends on the size of the information generated while crawling using R software. Because the process is still be done using manually flagged location it requires time to generate the information into the required format.

After information from social media has been converted into kml format, the next step is to import the information to my maps application and the information can be seen as shown in Figure 5. The information not only can be seen by the user but the user gives the facility to edit if the location or other information can be added to inform other users, the process shown in Figure 6. From the information that been added to the 'my maps' application user can choose wether the selection route from start to the assigned destination have blocked or other information gives by social media as identified in Figure 4.

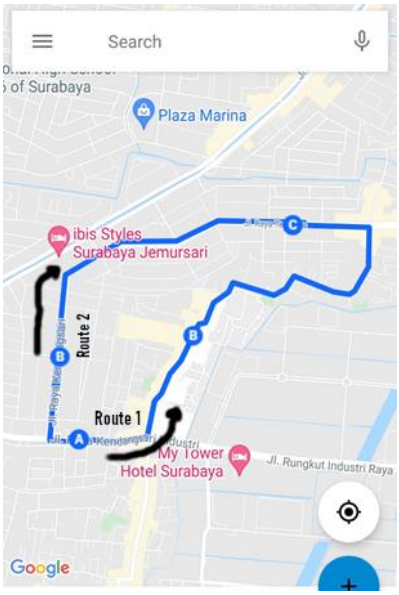

Figure 3. Maps of user route without any social media information and administrative borders

Figure 4. Maps user route of the location of the user option using social media information and administrative borders

Table 2. Social Media Information Result Extraction.

User id	Screen name	Result (text)	source
848614550	e100ss	09.14: 1. Raya Buduran Sidoarjo in the direction of Surabaya JAMMED, volume impact; 2. Raya Deket Lamongan, Gresik, is JAMMED. (hm)	Twitter Web App
848614550	e100ss	07.47: 1. The Waru Toll Road in the direction of the Satellite is JAMMED because a truck loading cement broke down in the left lane of KM 12,6. Vehicle queue since KM 15; 2. Aloha Sidoarjo roundabout in the direction of Surabaya is JAMMED, the volume is affected. (hm)	Twitter Web App
848614550	e100ss	21.28: 1. Yono Soewoyo Street in the direction of HR Muhammad JAMMED, the volume effect; 2. Klakah Lumajang in the direction of Surabaya is JAMMED. Because PPKM(hm)	Twitter Web App

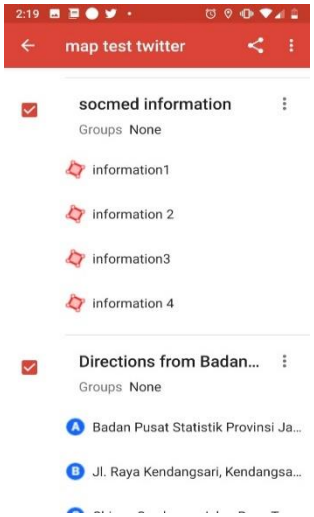

Figure 5. Social media imported information layer

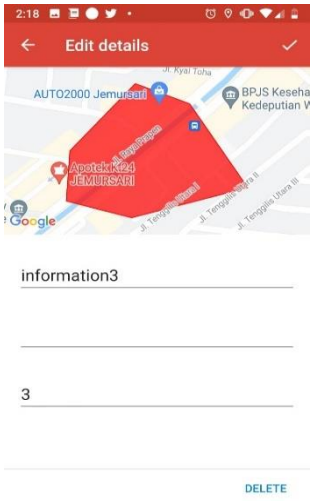

Figure 6. Social media information editing layer

Maps that are displayed to the user are synchronized with the maps in the web interface. The update process were be done while the user use the android application and the updated information can be imported as the user testing use the android application. To update the recent information the user must reload the maps from android devices. Table 3 showed the recommendation process on how the user finally selected the route that had minimum trouble while passing by.

Table 3. Recommendation final route survey officer

	Recommendation	Reason
By Location	Route 1	Based on location assignment
By Distance	Route 1	Distance calculation
By Path	Route 1	Based on calculation efectivity
By Socmed	Route 2	Because the results of crawling social media path to location, there is information there is a problem
Conclusion	Route 2	Taking into account the path and road information from social media

4.2. Discussion

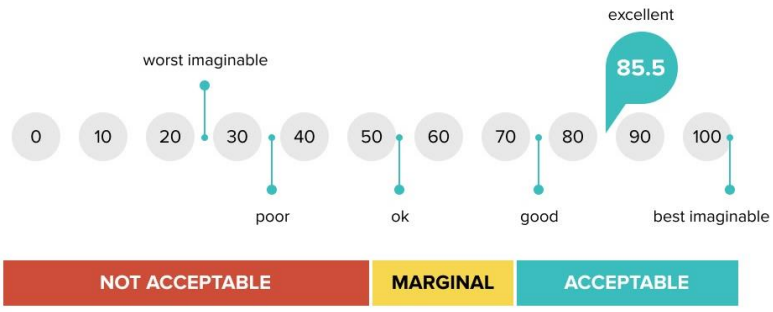

Figure 7. Acceptance of system usability score [20]

The testing of the system application were done after the selected user-tested and try to identify the location that had been assigned. The system testing in this research using System Usability Score (SUS). The SUS rate from the user gives a value of 65 on a scale of 100. Based on the SUS scale in Figure 7, the system is categorized as ok to good system. Although its have a range on 'ok' and 'good' range, but the value is almost on 'good' application range. The value were on the marginal range because some integration in this application its come from the information from social media that still done manually. The system can be leveraged to bigger system scope using scaling up the server that runs the extraction information from social media in automatic ways, then the information integrated into the custom maps application also information from social media can be seen by surveyor officer in realtime. Some users that had using the map said they had high acceptance because it can inform not only the route that will be passed by but also all the information from social media. The information that can be integrated in this application such as traffic jams, the restriction of survey location because of PPKM or other reasons, accident incidents and other information can be implemented and provided in the system.

5. Conclusion

The interpretation of the proposed mechanism can provide personalized travel route for a field survey officer. The survey officer can select the route based on the information that had been provided using the proposed mechanism. The information from social media source that gives location of traffic jam, restriction, and other information can help the surveyor to locate the road condition. The administrative border maps can be used by the field survey officer to identify the position of assigned survey location. Although the SUS score is 65 scale from 100 scales, the integration from the three different sources as proposed can be leveraged to more seamless. The proposed recommendation model mechanism is still limited to social media that provides public access. In the future, it can be improved by adding access from government sources such as care and protection applications where the status of the location zone whether it has high, medium, or low risk can be known. So that this information can be used by BPS survey field officers to anticipate possible risks of being exposed to Covid. Integration can also be carried out with related agencies related to disasters, or information from the police regarding accidents, road construction blocking, or other operations that cause field survey officers to consider while passing the road.

6. What to do next

In the future, the number of data sources can be increased to government-related sources. Thus increasing the synergy of performance from both the government and private sectors in supporting this route selection database. This route selection can be developed for other activities such as plans to use this route selection method for medical/health workers to get to the location faster, integration with the police to find out accidents, development of ambulance routes that are faster than integration with BNP and Polri. Firefighters and search and rescue teams can also use this model in advancing to the intended location. Integration can be done by combining the information available on social media of government agencies so that they can collaborate in providing the best route recommendations for users.

7. Acknowledgement

Thanks to statistics Jawa Timur that give access to use part of the administrative border data to be used in the testing of the application on this research. To all of my managers and colleagues in the IT division of statistics Jawa Timur that support to accomplishing this research.

Reference

[1] Fang J, Hu J, Shi X and Zhao L 2018 *Int. J. Disaster Risk Reduct.* Assessing disaster impacts and response using social media data in China: a case study of 2016 Wuhan rainstorm

[2] Fleischman R J, Lundquist M, Jui J, Newgard C D and Warden C 2008 Predicting Ambulance Time of Arrival to Emergency Department Using Global Positioning System and Google Maps pp 458–465

[3] Zhang Z, He Q, Gao J, and Ni M 2018 *Transp. Res. Part C* A deep learning approach for

detecting traffic accidents from social media data **86** pp 580–596
[4] Sujon M, Dai F and Asce M 2021 *J. Comput. Civ. Eng* Social Media Mining for Understanding Traffic Safety Culture in Washington State Using Twitter Data **35** pp 1–16
[5] Wang Y, He Z and Hu J 2020 *IEEE Trans. Intell. Transp. Syst.* Traffic Information Mining From Social Media Based on the MC-LSTM-Conv Model pp 1–13
[6] Putri E T and Ermawati M 2018 *International Conference on Informatics and Computational Sciences (ICICoS)* Real-time Detection of Data Completeness Degree for Traffic Simulation Using Text Similarity and Time Relevance of Data from Social **2** pp 1–6
[7] Pindarwati A and Wijayanto A W 2019 *Sixth Geoinformation Science Symposium SPIE* Personalized safety route recommender application using integrated urban crime hotspot data
[8] Al-bayari O 2020 *Int J Commun Syst.* An Internet and wireless networking-based water pipes web application for fault reporting pp 1–12
[9] Lavorgna L, Iaffaldano P, Abbadessa G, Lanzillo R and Esposito S 2021 *J. Neurol. Sci.* Disability assessment using Google Maps
[10] Rothfeld R, Straubinger A, Paul A and Antoniou C 2019 *J. Transp. Policy* Analysis of European airports ' access and egress travel times using Google Maps **81** pp 148–162
[11] Bibi S, Shah M A, Abbasi B Z and Hussain S 2018 *24th International Conference on Automation and Computing (ICAC)* A methodology to characterize and compute correlation between traffic congestion and health issues via social media pp 1–6
[12] Yang S and Hsu C 2015 *Comput. Electr. Eng.* A location-based services and Google maps-based information master system for tour guiding pp 1–19
[13] Smith M J 2008 *J Secure* Addressing the Security Needs of Women Passengers on Public Transport **1–2** pp 117–133
[14] Islam M, Shareef H and Mohamed A 2017 *J. PLoS One* Improved approach for electric vehicle rapid charging station placement and sizing using Google maps and binary lightning search algorithm pp 1–20
[15] Setiawan A and Sediyono E 2017 *International Conferences on Information Technology, Information Systems and Electrical Engineering (ICITISEE)* "A New Determination of Regional Area by Utilizing Rectangular Approach Method and Google Maps **2** pp 417–420
[16] Whitehead J, Smith M, Anderson Y, Zhang Y, Wu S, Maharaj S and Donnellan N 2021 *Int. J. Health Geogr.* Improving spatial data in health geographics : a practical approach for testing data to measure children ' s physical activity and food environments using Google Street View pp 1–15
[17] Badan-Informasi-Geospasial 2017 Satu data indonesia akses informasi geospatial terintegrasi https://tanahair.indonesia.go.id/portal-web (accessed Jul. 20, 2021).
[18] Google 2021 Google My Maps https://www.google.com/maps/d/u/0/?hl=en (accessed Jul. 20, 2021).
[19] Twitter 2021 twitter search https://twitter.com/home (accessed Jul. 20, 2021).
[20] Smyk A 2021 The System Usability Scale & How It's Used in UX https://xd.adobe.com/ideas/process/user-testing/sus-system-usability-scale-ux/ (accessed Oct. 07, 2021).