Paratransit Trip Data Collection System with Smartphone GPS and REST Web Service in Malang, Indonesia

Agi Putra Kharisma[1], Eriq Muh. Adams Jonemaro[2], Issa Arwani[3]
[1,2]Department of Informatics Engineering, Brawijaya University, Indonesia
[3] Department of Information System, Brawijaya University, Indonesia

Article Info	ABSTRACT
	Paratransit in Malang has its own irregular characteristics, same as in other cities in the world. It is tough to recognize the exact operational time, availability, and estimated travel time when using paratransit. The information about paratransit can be only obtained by making predictions due to its irregular characteristics. This research examined the importance of a system that can collect paratransit travel data as a basis for making predictions. Therefore, this research proposes the design of a paratransit travel data collection system by utilizing smartphone GPS and REST web service. To determine the position of the paratransit, the researchers used the concept of circle-based geofencing which is called pseudo-stop. The initial results obtained from the experiment indicated the influence of tracking intervals and radius of pseudo-stop on effectiveness and efficiency of the system. The shorter the tracking interval, the higher the tendency of the pseudo-stop radius value to decrease, resulting in decreased effectiveness and increased efficiency. Otherwise, the wider the tracking interval, the higher the tendency of the pseudo-stop radius value to be greater, resulting in increased effectiveness and decreased efficiency. In spite of periodically accessing location providers, the energy usage is quite low.

Corresponding Author:
Agi Putra Kharisma,
Department of Informatics Engineering,
Universitas Brawijaya,
8 Veteran Road, Malang, Jawa Timur 65145, Indonesia.
Email: agi@ub.ac.id

1. INTRODUCTION

Malang is the second largest city in East Java after Surabaya. With the advancement of information and communication technology (ICT), modern public transportations such as online taxis and *ojol* (online taxi bikes) are becoming very popular in Malang. Combining ICTs and informal modes of public transportation is predicted to be disruptive, although it has not been proven yet [1]. Meanwhile, more conventional modes of public transportation are still operating without adequate support from ICT. The researchers use the conventional term for transportation mode that is not supported by dedicated online location-aware mobile application. Paratransit is one of the conventional transportation modes which is still operating in Malang. Paratransit is a mode of public transportation owned and operated by private companies and individuals [2]. Paratransit has different names in various countries, for example, in Indonesia it is known as *angkutan kota*, in the Philippines it is known as *jeepney,* and in Thailand it is known as *tuk-tuk*. In Malang, paratransit is also known as *angkutan kota* (*angkot*) or *mikrolet*. However, the term *mikrolet* is commonly used. *Mikrolet* has similar characteristics with all of the paratransit in the world. *Mikrolet* does not have certain stops and operating hours. Passengers can get on or off from the *mikrolet* wherever they want because it does not have a specific stop location. *Mikrolet* has a fixed route but the route is not necessarily used because the drivers often take shortcuts or turn back before reaching the final destination.

In the past, *mikrolet* was one of the most popular public transportation modes in Malang. However, the current conditions have changed, *mikrolet* is not as popular anymore. Recently, the number of *mikrolet* is getting smaller and the operational hours are getting shorter. Similar conditions also occur in other cities. Many passengers have switched to online public transportation which is considered to have a better service. The main

factors of that issue are the uncertainty of waiting time, service time, and operating hours as well as number of vehicles in operation [3]. These problems make *mikrolet* as an unreliable transportation. Until now, the information about *mikrolet* is still very limited. Passengers have difficulty estimating trips with *mikrolet,* such as when the *mikrolet* arrives, how long it will take to travel, and whether there are still *mikrolet* in operation or not. In contrast, the scheduled transportation modes, such as Bus Rapid Transit (BRT) and Mass Rapid Transit (MRT), making predictions for trips of *mikrolet* become more complex. The implementation of ICT is expected to solve these obstacles.

The researchers want to overcome these problems with minimal changes to *mikrolet*. The system that the researchers are going to propose does not require change on how *mikrolet* operates. *Mikrolet* will continue to operate as usual, but the system will monitor and record the data on the trips of *mikrolet*. The *mikrolet* driver can drive as usual without much distraction from the proposed system. The researchers consider this to be beneficial for drivers, especially for drivers who are not familiar with technology such as the complex mobile apps. The collected data were used to create models to predict the information regarding *mikrolet,* such as its operating hours in specific days and the duration of travel on its specific routes. Certain machine learning techniques can be used to solve this problem by using predictive data analytics. Machine learning is the study of computer algorithms which improve automatically through experience [4]. One of the applications of machine learning techniques is for predictive data analytics. In order to do predictive data analytics, historical data are needed as a dataset to construct prediction models [5]. The process in establishing predictive data analytics solutions consists of various phases. One of the widely adopted process models is the Cross Industry Standard Process for Data Mining (CRISP-DM) [5][6]. The data must already exist before the processes of CRISP-DM can be run. The data act as a prerequisite before predictive data analytics can be performed. This research proposes a method for collecting data that can later be analyzed to produce a prediction.

In order to collect the data, the presence of *mikrolet* at certain times must be tracked. The method which is commonly used for vehicle tracking is GPS (Global Positioning System). Currently, GPS is widely available in most smartphones. GPS is tracking the data on smartphones and can be used for various purposes, such as to find accurate traffic information [7], to assess driving behavior [8], and to predict bus arrival [9]. The accuracy of GPS on smartphones is capable for recreational GPS classes where precision of the spatial data is not the highest priority [10], [11]. The researchers chose to use GPS on smartphones rather than dedicated GPS modules because it was easier to be found and maintained than specific GPS modules, although it may be having a lower accuracy [11]. Most of the vehicle tracking systems are built on a client-server architecture. Vehicle as a client to get location data, then the client sends the location data to the server and the server stores the location data sent from the client [7][9][12][13]. To conduct the data transmission from client to server, this research used REST (Representational State Transfer) based web service. REST based web service uses HTTP (Hypertext Transfer Protocol) because it is simpler, faster, and lighter alternative to SOAP (Simple Object Access Protocol) [14] and applicable for the case of this research. REST based services have been proposed for planning to travel with *mikrolet* [15][16]. The proposed innovation in this research is a design of *mikrolet* trip data collection system using the location data from smartphone GPS, which is then collected to an online server using REST based web service. The uniqueness of this research is that the data sent by the smartphone to the server is not the coordinates obtained from GPS; it is the location, which is called pseudo-stop. The researchers propose a pseudo-stop concept because *mikrolet* does not have a specific stop. Pseudo-stop can also be flexible to determine the checkpoint. With the pseudo-stop, the data stored on the server become orderly and do not require a lot of predictive data analytics.

2. METHODS

This research consisted of four stages as shown in Figure 1. In the first stage, the researcher developed API (Application Programming Interface) in the form of a REST web service for client applications as a destination for sending tracking data. Second, the researcher developed an Android-based application as a client application that will be deployed on smartphones in *mikrolet*. After the server and client were ready, the third step was to determine the pseudo-stop positions to simplify the tracking process. The final step, the researchers conducted an experiment and evaluated the results.
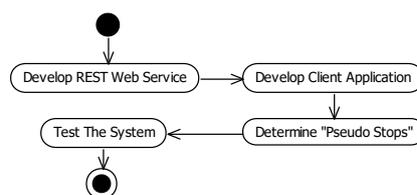

Figure 1. Steps in This Research

The architecture system can be seen in Figure 2. Most GPS modules on smartphones support A-GPS (Assisted GPS) which uses the help of mobile network cell sites, such as BTS (Base Transceiver Station) through mobile network resources. With A-GPS, the process of determining the location becomes faster and provides more energy efficiency because it might not have to deal directly with satellites. However, A-GPS can only work if it is connected to cellular network data, which results in additional costs. The application should be always connected with the internet in order to work properly. Internet connection holds two important roles, which are to obtain the latest location and to send data tracking to the REST web service server.
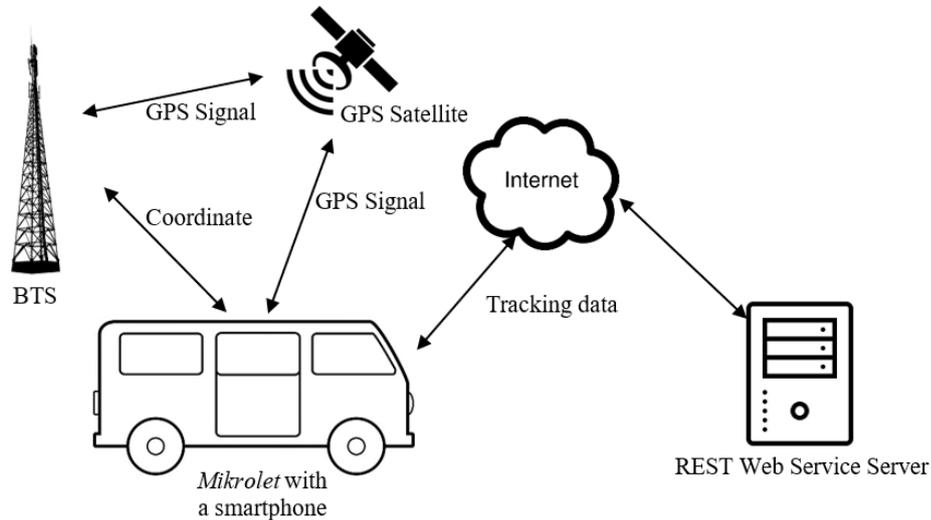

Figure 2. Proposed System Architecture

One of the main challenges in vehicle tracking systems is the battery life [7]. In order to save battery usage on the system, data transmission will be done when the position of the *mikrolet* is in the pseudo-stop area. When *mikrolet* is not in a pseudo-stop, there are no tracking data that needs to be reported. The pseudo-stop is similar to the circular geofence concept used in vehicle tracking [17]. Illustration of the *mikrolet* location data tracking process can be seen in Figure 3.
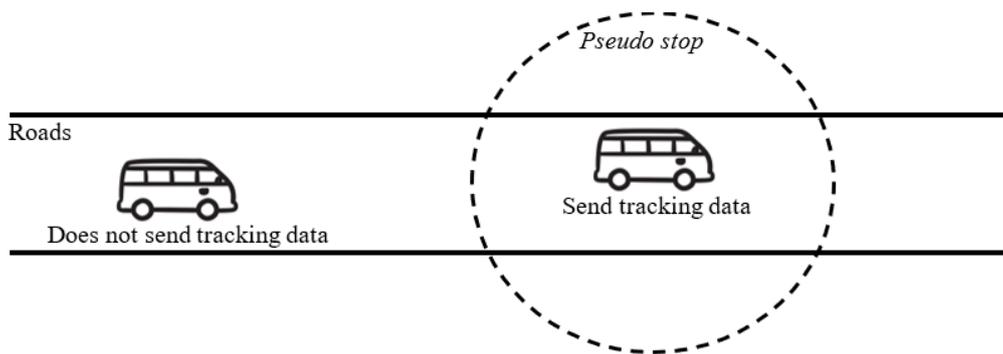

Figure 3. Illustration of Sending Tracking Data Process

2.1. REST Web Service Specification

The process of sending the *mikrolet* tracking data is done through the REST web service. In this research, the researchers applied Richardson Maturity Model level 2. Therefore, the HTTP method was used as a uniform interface [18]. There is only one entry point for sending tracking data, as the specifications can be seen in Table 1. At that entry point, HTTP POST was used because the new tracking data were added to the tracking service server database.

Table 1. The specification of REST Web Service

URL Path	HTTP Method	Parameters	Description

`/trip`	`POST`	*Mikrolet*, Driver, Latitude, Longitude, Time	The path `/trip` in URL serves to record the *mikrolet* trip data. This API is used periodically with HTTP POST method when *mikrolet* is within a *pseudo stop* location. The API consists of five parameters. The first parameter contains the name of *mikrolet* (*lyn*) along with its direction (inbound or outbound). *Lyn* is an abbreviation of the bus station that mikrolet pass through, for example a mikrolet with lyn AH means through Arjosari and Hamid Rusdi. To differentiate directions, we use the suffix -I for inbound and -O for outbound. The second parameter contains the driver ID. The third and fourth parameter contain latitude and longitude of visited *pseudo stop*. The last parameter contains the time when tracking data is sent to be recorded.
			
			 Example of HTTP POST request:
			 `POST /trip HTTP/1.1`
			 `Host: mikrolet-malang.com`
			 `Content-Type: application/x-www-form-urlencoded`
			 `Content-Length:117`
			
			 `mikrolet=AG-O&`
			 `driver=00001&`
			 `latitude=-7.968342802449274&`
			 `longitude=112.65269978823551&`
			 `time= 2020-04-01 15:04:05`

2.2. Client Application

Client application is an application that is installed on *mikrolet* driver's smartphone. The client application was created on Android SDK 29 and SDK 15 as minimum requirement and deployed on Android platform version 9 (Android Pie). To obtain the position of *mikrolet* in the form of latitude and longitude, the researchers used Android's fused location provider that combines data from GPS, cellular towers, and WiFi in order to gain optimal performance. Each version of Android may have different GPS management in order to optimize performance. The user of the client application is the *mikrolet* driver. The driver must activate the application and choose "start tracking" in order to report the tracking data. When the trip is over, the driver must stop the tracking process by clicking "stop tracking". The applied cases of client applications can be seen in Figure 4.
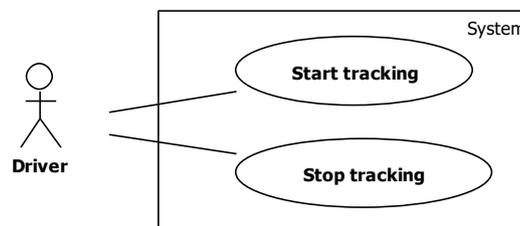

Figure 4. Client Application Use Cases

The process of collecting the current *mikrolet* location data is based on certain intervals. During the tracking process, the application will retrieve the current *mikrolet* location data and calculate whether the *mikrolet* is in the pseudo-stop area or not. An overview of the process of detecting the presence of *mikrolet* in a pseudo-stop can be seen in the following pseudo code:

```
Input: mikrolet current position p
Output: pseudo stop ps

Begin

let the result ps := null
let PseudoStops be the set of all pseudo stops
foreach pseudo stop pseudostop in PseudoStops begin
    if (distance between p and pseudostop <= radius of pseudostop) then
        ps := pseudostop
        break
    end if
end

End
```

As in the pseudocode above, if the distance between the coordinate of *mikrolet* location and the coordinate of the pseudo-stop is smaller than the radius of the pseudo-stop, then it can be concluded that the *mikrolet* is in that pseudo-stop. The coordinate of pseudo-stop is the center point of the circular pseudo-stop area, while the distance between two pseudo-stops is the distance between the two coordinates of the center of the pseudo-stop as shown in Figure 4. To calculate the distance between two pseudo-stop centers, the researchers used the haversine formula as in the following equation [19]:

$$d = 2r \, arcsin\sqrt{sin^2\left(\frac{\varphi_2 - \varphi_1}{2}\right) + cos(\varphi_1).cos(\varphi_2).sin^2\left(\frac{\lambda_2 - \lambda_1}{2}\right)} \tag{1}$$

Where d is distance between two points, r is radius of sphere (the radius of earth is approx. 6371 Km), φ_1 and φ_2 are the latitude of point 1 and latitude of point 2 (in radians), λ_1 and λ_2 are the longitude of point 1 and longitude of point 2 (in radians).

2.3. Pseudo Stop Placement

The researchers developed a method to help determine the precise location of pseudo-stops. In principle, the consideration in the placement and size (radius) of the pseudo-stop is affected by the speed of *mikrolet* and the tracking interval carried out by the client application. The method can be seen in the following formula:

$$2.r > \frac{1000.v}{3600}.t \tag{2}$$

$$d > 2.r \tag{3}$$

where r is radius of *pseudo stop*, v is speed of *mikrolet*, t is interval time of tracking and d is distance between two *pseudo stops*.

To produce optimal tracking data, each pseudo-stop must not intersect with each other and the distance between them must not be too far. A description of the relationship between the two pseudo-stops can be seen in Figure 5.
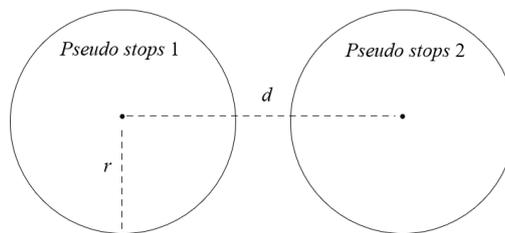

Figure 5. Relationship between two *pseudo stops* with distance (d) and radius (r)

In this research, the researchers determined that the maximum speed of an operating *mikrolet* was 50 km/h. This is based on the observations and government regulations. As for the tracking interval, the researchers used three variations which consist of five seconds, ten seconds, and 15 seconds. The values were obtained by applying these parameters as seen in Table 2.

Table 2. The Value of the Radius and the Distance Between Pseudo-stops

t (interval)	Minimum r (radius) value	r used in testing	Minimum d (distance between pseudo stops)	d used in testing/experiment
5 s	34.72 m	50 m	69,44 m	~110 m
10 s	69.44 m	100 m	138,89 m	~210 m
15 s	104.17 m	150 m	208,33 m	~310 m

As stated before, each pseudo-stop must not intersect with each other. If an intersection occurs, the tracking data become ambiguous. In order to secure the tracking data, the following implication should be always in order:

$$\forall x \in APS, \forall y \in APS, x \neq y \rightarrow distance\_of(x, y) > 2r \tag{4}$$

where *APS* is set of all *pseudo stops* traveled by specific *lyn* of *mikrolet* and *distance_of(x, y)* is distance between *pseudo stop x* and *pseudo stop y* calculated by the haversine formula and *r* is radius of *pseudo stops*.

By using the various formulas and rules mentioned above, the position of each pseudo-stop as test data can be determined. Visualization of the position of the pseudo-stops used as test data can be seen in Figure 6 to Figure 8. The coordinates of the pseudo-stops can be seen in Table 3 to Table 5.
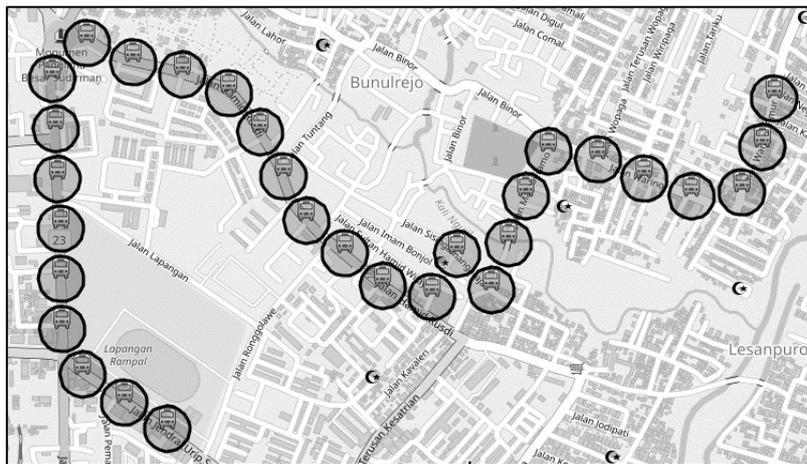

Figure 6. Visualization of Pseudo Stops for Five Second Tracking Interval

Figure 7. Visualization of Pseudo Stops for Ten Second Tracking Interval

Figure 8. Visualization of Pseudo Stops for 15 Second Tracking Interval

3. RESULTS AND DISCUSSION

3.1. Experimental Results

The researchers conducted an experiment by driving a car through the pseudo-stops listed in Table 3 to Table 5 and Figure 5 to Figure 7. The device used in this experiment was Xiaomi Mi A1 with the Android 9 operating system. The application used Android's ACCESS_COARSE_LOCATION permission to get an accurate location. In this research, the researcher made a single trip to get an initial overview of the tracking system's performance. During the trip, the researcher carried out three tracking intervals at once to obtain the most homogeneous conditions. The number of reports for each pseudo-stop at various tracking intervals can be seen in Table 3 to Table 5.

Table 3. The List of Pseudo Stops with Five Second Tracking Interval

No.	Coordinate (Latitude, Longitude)	Distance from the Nearest Pseudo Stop (meter)	Number of Recorded Reports
1	-7.968395928681576,112.65268905939945	110.61	0
2	-7.969362824904677,112.65245302500614	110.61	2
3	-7.97026596766513,112.65204532923588	110.00	3
4	-7.970340344038929,112.65104754748234	110.18	7
5	-7.969989712444104,112.6501034099091	111.03	5
6	-7.969607204907132,112.64918073000797	110.14	3
7	-7.969447826661214,112.64819367709049	110.13	8
8	-7.970340344038929,112.64774306597599	110.95	0
9	-7.971275360156284,112.64739974322208	110.63	1
10	-7.972210374136991,112.64705642046818	110.63	8
11	-7.971487863521349,112.64636977496036	110.33	0
12	-7.972348501021161,112.64586551966556	110.64	2
13	-7.9721253729543164,112.64488919558414	110.34	2
14	-7.971477238355718,112.64412744822391	110.59	2
15	-7.970829102730424,112.64336570086368	110.59	2
16	-7.969947211624319,112.64290436091312	110.44	1
17	-7.969065318617771,112.64244302096256	110.44	2
18	-7.968289676210076,112.6418207484711	110.15	2
19	-7.967917792342628,112.64088733973392	110.79	2
20	-7.967673411331478,112.6399110156525	110.89	1
21	-7.967290901629859,112.63898833575138	110.15	4
22	-7.968002794399256,112.6382909614075	110.29	2
23	-7.969001567242445,112.63836606325992	111.37	1
24	-7.97000033764836,112.6383982497681	111.11	2
25	-7.9709884804383835,112.63846262278446	110.10	1
26	-7.971987245994793,112.63847335162052	111.06	1
27	-7.972986009113053,112.63849480929264	111.08	14
28	-7.973889143875874,112.6389025050629	110.00	2
29	-7.974452273953969,112.63972862543949	110.44	1
30	-7.9749729029590455,112.64058693232425	110.84	1
	Total number of recorded reports		82

Table 4. The List of Pseudo Stops with Ten Second Tracking Interval

No.	Coordinate (Latitude, Longitude)	Distance from the Nearest Pseudo Stop (meter)	Number of Recorded Reports
1	-7.968342802449274,112.65269978823551	210.72	8
2	-7.970138465278598,112.65208824458011	210.72	3
3	-7.970021588056044,112.65018924059757	209.52	3
4	-7.969320324019792,112.64841898264774	209.96	8
5	-7.970967230080913,112.64748557391056	210.01	1
6	-7.972135998103116,112.64598353686222	210.35	5
7	-7.9714984886867,112.64418109240421	210.76	2
8	-7.970085339272483,112.64291508974918	210.07	1
9	-7.968427804417647,112.64199240984806	210.46	1
10	-7.967737162913668,112.64021142306217	210.62	2
11	-7.967769038701049,112.63830169024357	210.33	3
12	-7.969660330975305,112.63837679209598	210.46	2
13	-7.971551614509335,112.6384518939484	210.46	1
14	-7.973442889301101,112.63851626696476	210.42	8
15	-7.974601025165897,112.64002903284916	210.56	2
	Total number of recorded reports		50

Table 5. The List of Pseudo Stops with 15 Second Tracking Interval

No.	Coordinate (Latitude, Longitude)	Distance from nearest *pseudo stop* (meter)	Number of recorded reports
1	-7.968300301458465,112.65272124590763	309.97	7
2	-7.970244717270129,112.65070422472843	309.97	3
3	-7.969936586418677,112.64791472735294	309.09	6
4	-7.972040371753983,112.64605863871463	310.64	6
5	-7.970903479001861,112.64348371806034	310.45	1
6	-7.968523431610965,112.64202459635624	309.61	2
7	-7.967503407063774,112.6394067603577	309.7	4
8	-7.970106589675754,112.63837679209598	310.89	2
9	-7.9728903829626026,112.63848408045658	309.77	6
10	-7.9749197775807215,112.6404152709473	310.08	3
	Total number of recorded reports		40

Next, the recorded tracking data were accumulated based on the number of pseudo-stops that should be passed, the number of recorded pseudo-stops, the number of pseudo-stops that should be passed but not recorded, and the amount of data recorded as in Table 6. To calculate the effectiveness and efficiency, the following equations are used.

$$Effectiveness = \frac{Number\ of\ recorded\ pseudo\ stops}{Number\ of\ pseudo\ stops} \times 100\% \qquad (5)$$

$$Efficiency = \frac{Number\ of\ pseudo\ stops}{Total\ number\ of\ recorded\ reports} \times 100\% \qquad (6)$$

Table 6. Accumulated Experimentation Data

Tracking interval	Number of pseudo-stops	Number of recorded pseudo-stops	Number of unrecorded pseudo-stops	Number of recorded reports	Effectiveness	Efficiency
5	30	27	3	82	90	33
10	15	15	0	50	100	30
15	10	10	0	40	100	25

Apart from that, the researchers also checked the energy consumption required by the application. The researchers used the energy profiler in Android Studio to examine the predicted energy usage within one minute (60 seconds) for each tracking interval. The energy profiler in Android Studio does not measure the exact values, but only approximate results that are displayed in graphical form with certain levels, such as light and medium. The results can be seen in Figure 9 to Figure 11.
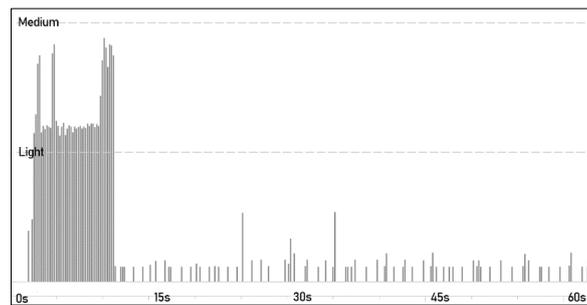

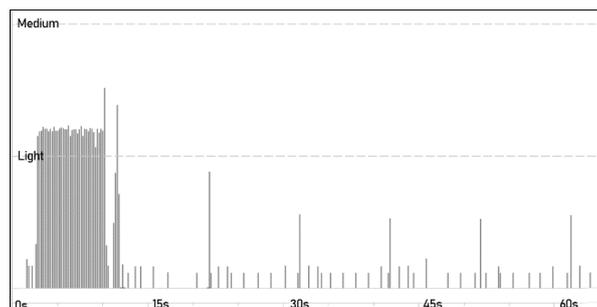Figure 9. Energy Usage for Five Second Tracking Interval

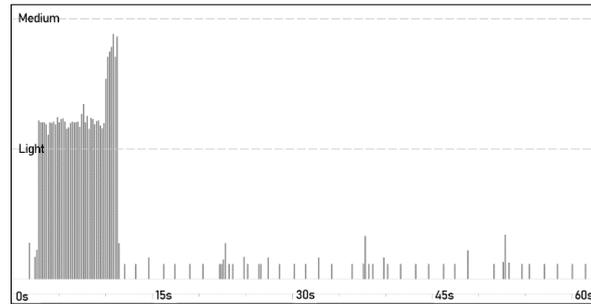Figure 10. Energy Usage for Ten Second Tracking Interval

Figure 11. Energy Usage for 15 Second Tracking Interval

3.2. Discussion

The system is greatly influenced by the concept of geofencing for tracking moving objects [17][20][21] but with different cases and some adjustments which only use circular geofence as pseudo-stop. Circular geofences have been used to monitor vehicles on predetermined routes [17], but the method for determining the radius or diameter of the geofences is not described. As a comparison, in this study the researchers offered a method to determine the size of a circular geofence radius and conduct the experiments in the city of Malang, Indonesia to determine the efficiency, effectiveness, and energy usage which has never been done in other studies.

Based on the results of the experiment, the smaller the value of the tracking interval, the higher the efficiency but the lower the effectiveness. Otherwise, the greater the value of the recording interval results in the lower efficiency, the higher the effectiveness, as shown in Figure 12.
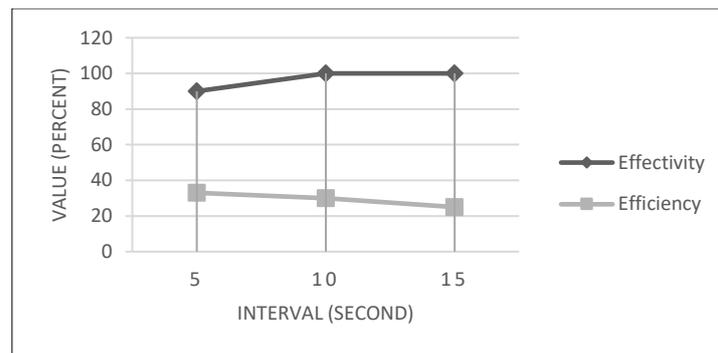

Figure 12. Trends in Effectiveness and Efficiency Based on Tracking Intervals

In experiments with five second tracking intervals that can be seen in Table 3, there were some pseudo-stops that were not recorded; therefore the effectiveness did not reach 100%. This phenomenon occurred due to the function of the Android operating system in saving battery usage. When a client application requests a current location, the Android operating system might send an expired location. In other cases, the programmed tracking interval in client application can also be overridden by the Android operating system. For example, when the client application runs with five second tracking interval, in fact, it may be higher or lower than five seconds due to overrides made by the Android operating system. If the tracking interval is higher than it should be, the efficiency can be lower, whereas if the tracking interval is less than it should be, the effectiveness has the potential to decrease. These conditions are beyond the application's control because location services are dynamically controlled by the Android operating system by considering the efficiency of battery usage and other factors. In experiments with five second, ten second, and 15 second tracking intervals that can be seen in Table 3 to Table 5, there were also pseudo-stops that were recorded repeatedly up to 14 times, so the efficiency did not reach 100%. This can happen when *mikrolet* runs at low speed or stops.

Based on experimental results in Table 6, effectiveness shows a promising value (~ 90%), while efficiency has a low value (~ 30%), so they still need to be improved. To increase effectiveness at relatively small tracking intervals, it can be done by increasing the size of the geofence radius, but this method can also result in lower efficiency. In addition, if the pseudo-stop radius is too large, it will be difficult to be applied on winding roads, turnarounds, and crossing roads, because the possibility of overlapped pseudo-stops becomes greater. Overlap between pseudo-stops can result in ambiguous recording. To improve efficiency, it can be done by creating a filter in the client application or on the REST web service server to check the presence of data before storing them, so that the recorded pseudo-stops are not recorded repeatedly during a trip.

In the results of energy use (Figure 9 to Figure 11), it appears that the smaller the tracking interval, the more the energy consumption will be. This can be seen from the density of the energy consumption lines. The spike in energy usage that occurs in about the first ten seconds is the energy used to start the application. However, the use of this energy is still relatively low because it has not yet reached the medium level. After that, energy consumption tends to be stable with quite low usage. In general, the difference in energy consumption between the five second, ten second, 15 second intervals may not be that significant due to many influencing factors. One of the factors that plays a big role in energy consumption when accessing location is the presence of a fused location provider on Android that combines various sensors, resulting in a much better energy use efficiency.

3.3. Research Limitations

The experiments conducted in this research still have many limitations. Variations on the conditions in experiments have not been carried out, for example, experiments in areas with limited cellular signals and areas with heavy traffic. This research only used one type of smartphone to conduct all experiments, so that there was no comparison between one type of smartphone with another.

4. CONCLUSION AND FUTURE WORK

4.1. Conclusion

This research has proven that paratransit trip data collection system can be implemented using GPS-based tracking on smartphones by utilizing the concept we propose, namely pseudo-stop, as well as data communication via the REST web service. Initial results obtained from the experiments that have been carried out indicate the influence of the tracking interval and the pseudo-stop radius on the effectiveness and efficiency. Furthermore, the shorter the tracking interval, the higher the tendency of the pseudo-stop radius value to decrease, potentially resulting in decreased effectiveness and increased efficiency. This is in accordance with the results of previous experiments that have been conducted. Conversely, the wider the tracking interval, the higher the tendency of the pseudo-stop radius value to be greater, so that it can result in increased effectiveness and decreased efficiency. This is also in accordance with the results of previous experiments that have been conducted. In addition, in spite of periodically accessing location providers, energy usage is quite low.

4.2. Future Work

To produce more convincing generalizations as the conclusions, more varied experiments are needed. Since there is no guarantee that the value of effectiveness will always be 100%, it is necessary to develop an error correction mechanism for error data that are interrupted due to missed pseudo-stops. In addition, because there is no guarantee that the efficiency value will always be 100%, tracking data need to be filtered, whether filtered by a smartphone before being sent via the REST web service or filtered by the REST web service server after the data are sent. With these steps, the collection of travel data can be clean.

REFERENCES

[1] R. M. Medeiros, F. Duarte, F. Achmad, and A. Jalali, "Merging ICT and informal transport in Jakarta's ojek system," *Transp. Plan. Technol.*, vol. 41, no. 3, pp. 336–352, 2018, doi: 10.1080/03081060.2018.1435465.

[2] T. B. Joewono and H. Kubota, "User satisfaction with paratransit in competition with motorization in indonesia: Anticipation of future implications," *Transportation (Amst).*, vol. 34, no. 3, pp. 337–354, 2007, doi: 10.1007/s11116-007-9119-7.

[3] J. H. Frans, J. J. S. Pah, and M. G. A. Ikun, "Perpindahan Moda Angkutan Umum Ke Angkutan Pribadi Di Kota Kupang," *J. Tek. Sipil*, vol. 6, no. 2, pp. 151–164, 2017.

[4] T. M. Mitchell, *Machine Learning*, 1st ed. USA: McGraw-Hill, Inc., 1997.

[5] J. D. Kelleher, B. Mac Namee, and A. D'Arcy, *Fundamentals of Machine Learning for Predictive Data Analytics: Algorithms, Worked Examples, and Case Studies*. The MIT Press, 2015.

[6] R. Wirth, "CRISP-DM: Towards a standard process model for data mining," in *Proceedings of the Fourth International Conference on the Practical Application of Knowledge Discovery and Data Mining*, 2000, pp. 29–39.

[7] J. C. Herrera, D. B. Work, R. Herring, X. (Jeff) Ban, Q. Jacobson, and A. M. Bayen, "Evaluation of traffic data obtained via GPS-enabled mobile phones: The Mobile Century field experiment," *Transp. Res. Part C Emerg. Technol.*, vol. 18, no. 4, pp. 568–583, 2010, doi: 10.1016/j.trc.2009.10.006.

[8] Q. C. Sun, R. Odolinski, J. C. Xia, J. Foster, T. Falkmer, and H. Lee, "Validating the efficacy of GPS tracking vehicle movement for driving behaviour assessment," *Travel Behav. Soc.*, vol. 6, pp. 32–43, 2017, doi: 10.1016/j.tbs.2016.05.001.

[9] D. Bhardwaj, S. Daphal, and A. Nerkar, "Predicting Bus Arrival Time with GPS on Android Application," *Int. J. Sci. Res.*, vol. 5, no. 1, pp. 122–125, 2016, doi: 10.21275/v5i1.nov152627.

[10] K. Merry and P. Bettinger, "Smartphone GPS accuracy study in an urban environment," *PLoS One*, vol. 14, no. 7, pp. 1–19, 2019, doi: 10.1371/journal.pone.0219890.

[11] R. Goodspeed *et al.*, "Comparing the data quality of global positioning system devices and mobile phones for assessing relationships between place, mobility, and health: Field study," *JMIR mHealth uHealth*, vol. 6, no. 8, 2018, doi: 10.2196/mhealth.9771.

[12] A. Shingare, A. Pendole, N. Chaudhari, P. Deshpande, and S. Sonavane, "GPS supported city bus tracking & smart ticketing system," *Proc. 2015 Int. Conf. Green Comput. Internet Things, ICGCIoT 2015*, pp. 93–98, 2016, doi: 10.1109/ICGCIoT.2015.7380436.

[13] S. Bind and P. R. Solanki, "GPS Tracking System for Public Transport," *Int. Res. J. Eng. Technol.*, vol. 4, no. 6, pp. 2600–2603, 2017, [Online]. Available: https://irjet.net/archives/V4/i6/IRJET-V4I6504.pdf.

[14] J. Tihomirovs and J. Grabis, "Comparison of SOAP and REST Based Web Services Using Software Evaluation Metrics," *Inf. Technol. Manag. Sci.*, vol. 19, no. 1, pp. 92–97, 2017, doi: 10.1515/itms-2016-0017.

[15] A. P. Kharisma and A. Pinandito, "Design of REST API for Local Public Transportation Information Services in Malang City," *J. Inf. Technol. Comput. Sci.*, vol. 2, no. 2, pp. 92–102, 2017, doi: 10.25126/jitecs.20172226.

[16] A. P. Kharisma and A. Pinandito, "Prototype of REST Location-based Service for Travel Planning by Local Public Transportation (Mikrolet) in Malang City," *J. Inf. Technol. Comput. Sci.*, vol. 4, no. 3, p. 253, 2019, doi: 10.25126/jitecs.201943131.

[17] F. Reclus and K. Drouard, "Geofencing for fleet & freight management," *2009 9th Int. Conf. Intell. Transp. Syst. Telecommun. ITST 2009*, pp. 353–356, 2009, doi: 10.1109/ITST.2009.5399328.

[18] E. Wilde and C. Pautasso, Eds., *REST: From Research to Practice*. Springer, 2011.

[19] W. M. Smart, *Textbook on Spherical Astronomy*, 6th ed. Cambridge University Press, 1977.

[20] L. S. Humphries and H.-J. Ngo, "Method and system for tracked device location and route adherence via geofencing," US7164986B2, 2007.

[21] W. Duvall, "Geofencing moveable vehicles," EP1433035A2, 2003.

BIOGRAPHY OF AUTHORS

Agi Putra Kharisma is a lecturer in the Informatics Engineering study program at the Department of Informatics Engineering, Faculty of Computer Science, Brawijaya University. He specializes in software engineering, specifically model-driven software engineering and service-oriented architecture. He is also the coordinator of lecturers in the fields of multimedia, games, and mobile technology in the department of Informatics Engineering, Faculty of Computer Science, Brawijaya University.

Eriq Muh. Adams Jonemaro is a lecturer in the Informatics Engineering study program at the Department of Informatics Engineering, Faculty of Computer Science, Brawijaya University. He specializes in game technology. He is also the head of learning laboratory in Faculty of Computer Science, Brawijaya University.

Issa Arwani is a lecturer in the Information Technology study program at the Department of Information System, Faculty of Computer Science, Brawijaya University. He specializes in wireless sensor network technology. He is also the head of Information Technology study program at the Department of Information System, Faculty of Computer Science, Brawijaya University.