Advances in Dynamical Systems and Applications.
ISSN 0973-5321, Volume 16, Number 2, (2021) pp. 1321-1331

https://www.ripublication.com/adsa.htm

Accuracy improvement technique of Big Data based LBS system: using recommendation system

Changbae Mun

Hanyang Cyber University, Seoul, South Korea

ABSTRACT

As the LBS service gradually develops, this technology is being used in many areas of life. In particular, the LBS application is a service that provides various information related to a user's location quickly through wireless communication to a moving user, and efficiently provides information about a moving person. From this point of view, since it is a service that provides information to a moving object, it is necessary to recommend a more accurate route. This study designs a recommendation system for LBS applications and explores ways to provide more accurate information.

Keywords - LBS, Big Data, System architecture, Intelligent platform

I. INTRODUCTION

The Location based service (LBS) is a service that provides location information obtained by using various communication networks to users [1]. This service is gradually developing in the form of various services as the information collected by the LBS application is gradually accumulated, and is expected to reach 96 billion in 2021 [2]. In particular, various information recommendation functions are provided for the convenience of providing information to users in the commerce industry [3]. With the development of travel-related applications, there are cases in which restaurant review information is provided along with location information to activate a new urban landscape, and convergence and linkage with various information systems is frequent. Recently, as the demand for electric vehicles has rapidly increased, an LBS system has been proposed that selects a place to charge electricity and provides information about the surrounding commercial area and customized advertisements and information during the charging standby time [4]. It is also being continuously introduced in the field of public infrastructure. For example, for the purpose of revitalizing traditional markets in Korea, marketing activities of sellers are supported through LBS and various information such as product information is provided to users [5]. In particular, the development potential of LBS has been confirmed in mobile services, and it appears

that LBS provides the benefit of providing appropriate information to users in a context-based manner and statistically affects the perceived value. This technology can reliably provide users with the information they need. So this research provide an architecture that LBS system need to consist in their core engine[6]. Summarizing the above, the LBS system is being developed in the direction of providing accurate information to the user of the application and system modelling the movement information required for the user. As a result, LBS is in the process of increasing user satisfaction. In this study, we tried to find various methods that can provide users with accurate routes for movement information. In order to analyze the route information of the area selected by the user, the best route is determined that the residents of this area have relatively high accuracy information, and the user's movement data is used through data analysis to apply the method to LBS.

II. STUDY OBJECTIVE

In the case of the route-based LBS system among previous studies, information guidance services using low-power Bluetooth beacons are provided in indoor spaces where GPS is not supported [7]. Regarding external movement information, the previous study of this study showed that movement information created when a user moves through a fixed external path is primarily analyzed in the information processing server. This result corresponds to a technique for statistically extracting the routes traveled by many users for the same route. Secondarily, for specific area information, if there is sales information of a store related to the route, the user compares it with the past movement information and sends the relevant information so that the user can move in a similar way [8]. An analysis of the products of companies utilizing LBS and information of local companies was also performed. This study analyzed the similarity between information and compared the high similarity through the collaborative filtering technique. If the similarity is high, the probability of user visit is high. In this case, recommended information was provided to the user, and as a result of the experiment, it was confirmed that the probability of the user's visit was high [8]. A study on the personalization system proposed a personalized function by applying a recommendation system to LBS using collaborative filtering technology. Based on the current location, this study analyzes the contents open to all users among major places within a distance adjacent to the user by preference unit. As a result, it provides a list of contents that are expected to have a high probability of users visiting [9]. Among them, the study of [8] focused on providing personalized characteristics. In navigation, which is used on a daily basis, it is often not easy to check accurate movement information depending on the weather. In addition, in this system, when there is road construction in the route, it is difficult to receive information about the corrected information in the navigation system. Also, from the user's point of view, in the case of the first route the user travels, the travel time varies according to the accuracy of weather and map information and the complexity of the terrain.

In order to solve such difficulties in the insufficient routing information, this paper analyzed the user's movement information and deals with a method of guiding through

the recommendation function. To this end, this technology links with the location processing platform of the portal map information service, provides dynamically refined information by utilizing each user's location information, and provides a more accurate actual service based on the user's surrounding information. do. , In this paper, the movement information of local residents is collected in the form of big data and the most efficient route information is analyzed. Design the structure of a statistical analysis model based on the map information of navigation. As a result, the central server analyzes each user's movement information in real time, so that the system provides information in a way that guides the user to the shortest distance. As users will be provided with accurate information more conveniently through this service, the LBS service will evolve into a platform industry related to web services.

III. METHOD

3.1. System design

The In The structural design of the system is constructed from the perspective of the overall schema of the system for path information analysis. In this schema, the collection and analysis function of the user's movement information data is used first. This result is used to configure mobile applications related to user notifications.

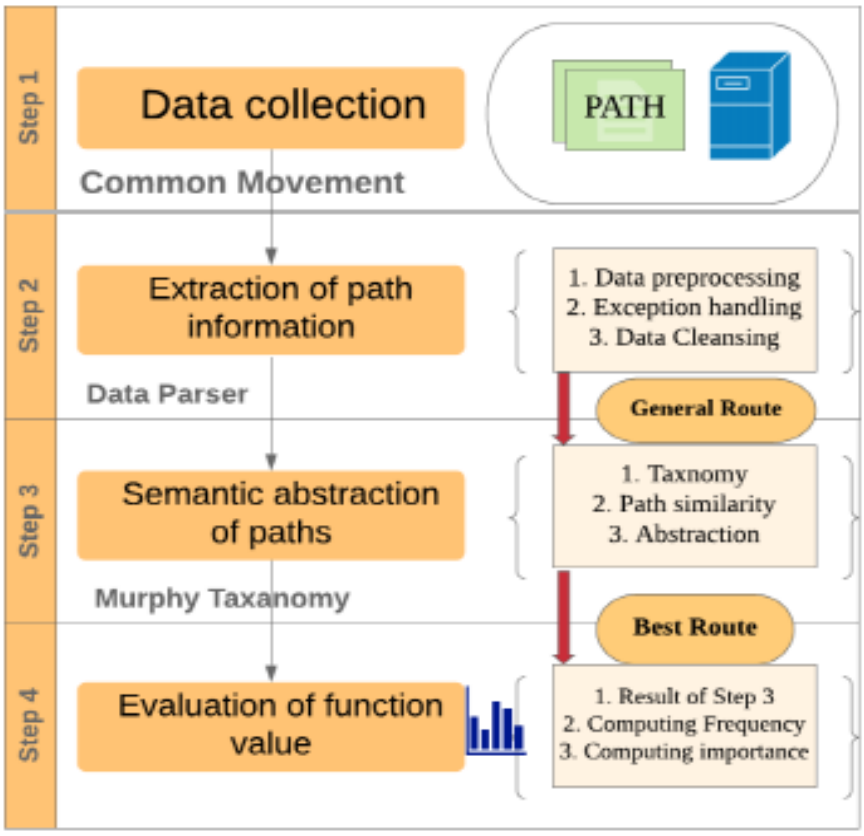

Fig. 1. LBS-based information processing [15]

Unlike a generally used navigation system, a mobile information system based on user big data has a function similar to that of a recommendation system. In this recommendation system, the user's information is accumulated in the process of moving the user, and the route information is provided to others. In addition, as the opposite process of the process, the user is consumed and provided with information. The LBS system consists of a system that purifies and analyzes these two types of information. In this process, user data is continuously collected and the user's repetitive route information is processed. This study introduces a big data framework to implement data real-time streaming flow on a distributed system. The big data framework presented in the previous study of this study consisted of a Restful API module built with the Spring open source framework for application users, a web service linkage module, and a relational database [15]. Information interworking for mobile information analysis is supported by Rest API. This analysis server refines the route information and continuously communicates to the server as individual clients.

In this way, the API module operates as a design structure to manage the communication process, and each transmission function is configured by dividing the main menu of the interworking function. This function processes user requests in a system to which distributed processing is applied, and primarily purifies the movement information accumulated data of individual users. And among them, the method of checking highly useful information and storing it in the database was used. In terms of real-time In terms of real-time analysis function, efficiency is improved through the link between Hadoop and Spark, which are open sources of the Apache project. Based on YARN provided by the Hadoop ecosystem, Spark is driven and the recommended information data, which is important in real-time, is transferred to Apache Spark for processing. and configure a personalized recommendation module to provide route information in accordance with an individual's characteristic method such as a user's movement pattern. From the user's point of view, as shown in Figure 1, when the system user inputs route information into the mobile application and starts walking, the system starts collecting information. At this time, the information processing server is set to transmit the user's movement information to the server at a predetermined period. When this time is set as small as 5 seconds or less, the system load used for exception handling becomes significant. In particular, in the analysis process for information recommendation, it is important to find the optimal collection period because each turning point is composed of nodes. In order to check the route information, the recommendation module of this system needs to verify route information that has been collected and refined for a long time on the route. This information needs to be verified and applied to real user data in the region. In this aspect, there is a process to support the inspection system of local users by applying a reward system such as functional benefits or points of the system itself. The system can use this method to refine the values to be more accurate and to keep the data up-to-date.

3.2 Architecture with Web System

When This study analyzed the movement data by referring to the system design of previous studies for the introduction of the data analysis system. The system constitutes a real-time stream for full analysis, excluding personal information related to privacy violations from among the data collected by each user. The data schema of the analysis system is divided in the receiving module of the data processing system. It can be implemented as a route guidance function through a real-time and batch analysis function of the collected user information and a notification of a smartphone application. The setting of providing an optimal shortcut requires that residents become information providers about the main crossroads and fast routes of the starting point and destination of a specific route. For this role, a reward system for the user is applied to strengthen the information provision activity of this user and prepare more accurate route information. In the overall structure, heterogeneous systems can access route information data through each data node in the system network configured with Hadoop through HDFS, an open source, and support efficient parallel processing in a clustered system environment using Map Reduce, a distributed processing technology.

It is designed to provide real-time streaming data to each user faster and without loss by introducing the scheduling of the detailed data node unit into the data processing architecture. In summary, SPARK is introduced to analyze the user's movement information in terms of the stability of collection. In addition, this system applies KAFKA for distributed delivery of data sets and commonly applies the FLUME module for the purpose of performing individual collection tasks in Hadoop. And this system introduces KAFKA, an open source distributed messaging system technology to expand the distribution of data. We also needed FLUME for batch jobs to be used in Hadoop, so we drop this open source according to common requirements. The FLUME module of this system provides the function of processing big data corresponding to movement information faster and loading it into logs, etc. in the linked server that provides real-time streaming service. This module is used to transfer and load large log information from heterogeneous collection equipment. A process of efficiently processing large amounts of data delivered by real-time streaming is used. In particular, this process is characterized by high stability at the level of the system structural layer. This process is also effective in collecting unstructured information such as mobile information, and it can collect all types of structured/unstructured data including server usage information in the system, and compatibility with heterogeneous systems through multiple protocols based on extensibility. It has a high advantage. In this process, the system collects API linkage information from the controller by modifying only its own setting properties for structural design. The information delivery module of the system is designed in a way that logs are input to HDFS for batch processing. In relation to route information processing, route information produced by mobile information providers is applied to Kafka, and Stream is configured using StreamingContext. In this process, SPARK structure streaming technology is applied to implement functions on HDFS, collect movement information, and analyze it structurally in a table format through real-time streaming processing.

3.3 Complementing the stability of mobile information data

For the stability of data linkage, it is composed of a multi-instance structure suitable for big data processing. To this end, a module that improves the performance by applying the dynamic distribution function in the system is loaded. This process ensures system stability and ensures data integrity by ensuring message delivery. And in terms of system operation, it is supported to change the connection module quickly to the environment change situation. From this point of view, support for a seamless environment for various interfaces is supported by the linkage adapter, and support for the standardized framework-based linkage environment configuration is applied to be applied to web services. In terms of stability within the module, it supports various types of data types and supports diversity of interfaces through integrated routing function between heterogeneous web services. As an essential function among data linkage functions, real-time data generated by applications constituting the mobile information service is collected at specific intervals to verify and apply the consistency of information. It provides methods such as periodic collection through batch jobs for information collection, and consistency checks on data refined through an independently configured analysis system are conducted in batch jobs.

In addition to the stability of server functional data, it is necessary to review measures to increase the stability and consistency of the collected data itself. The collection server provides route information by analyzing the user's location in units of a certain time and analyzing user data on the progress direction and the history of the current route. Depending on this case, the reliability of the previously stored data may be affected, so it is necessary to continuously classify the exception cases separately through probabilistic calculations. In addition to the normalized path, exceptional movement information is also continuously accumulated in the server, thereby accumulating know-how on the system for exception handling. The error is compensated for by separately providing the user with the best route and the choice of movement through the analysis result of only this exception information. Various situations may occur on a general road, and for various paths provided in each situation, users may select and use the path with the highest stability at the present time.

3.4 Analysis data system processing logic

The purification process and flow of system data starts with the collection of data generated by the information provider corresponding to the system user. Based on the initially created route data, information of various users traveling to the same starting point and destination is additionally collected. The system configured in conjunction with the data analysis module has the following three representative functions to guide more accurate movement information. This classification supplemented the research methodology used in previous studies [16].

Function 1: Judging users' common movement paths as optimal shortcuts

Function 2: Classify unusual items in the user's moving path and extract exception information

Function 3: Validate temporary exception information according to a specific individual's exception case or situation and present the result to the user

First, Function 1 is to analyze the detailed route common to residents within a specific movement route, based on previous studies [13]. This corresponds to a route traveled by local residents who are well aware of the topography and characteristics of a specific route, and users empirically verify that it is the optimal route. Therefore, this route becomes the first guide reference information for route guidance and is the best information unless a specific exception of function 3 occurs. And, the destination of this route is highly likely to be close to the starting point of public transportation such as village buses and subway stations, and if the slope or topography is not complicated, it is highly likely to be the shortest route on a straight line. In order to statistically analyze this route, all the movement routes of residents are collected in specific time units such as 24 hours and 48 hours. The system analyzed the movement path trend from this information. As a result of this process, the optimal movement route that the local residents used empirically was converted into data. In order to increase the accuracy of this function, various exceptions such as weather are applied in the data linkage system. Of course, it is necessary to statistically classify exceptions, but if you select and start the process of creating exception data at the beginning of system use, you can further reduce errors. Fig. 2 shows the process of the data refining procedure.

The information processing server of the system is designed to transmit the user's current location to the server in a reference time unit. In this system, coordinate information for the user's movement information is stored in the server in real time. With respect to the characteristics of the user's movement, even a local resident may have several routes for moving to the same destination. The user can be referred for moving by receiving the statistically analyzed latest moving information. In this process, it is possible to refine the similar route information about the movement information of all users.

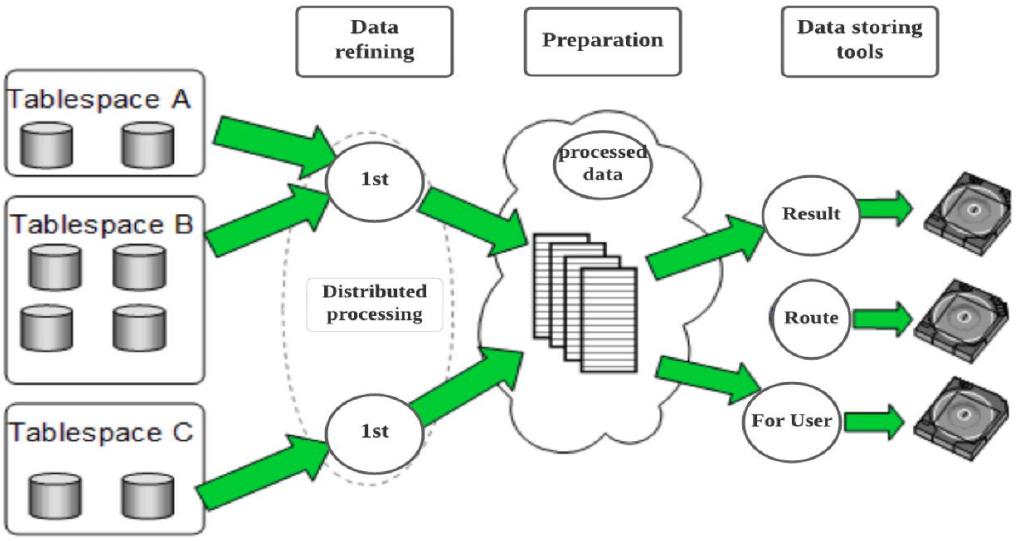

Fig. 2. Flow of data analysis procedure

IV. EXPERIMENT

To verify the effectiveness of this system, a mobile information simulation experiment was designed and implemented for a specific area. Exceptional information that can verify the characteristic functions of this system is set in the path set for the experiment. 150 cases of route information that can be used as movement information of actual local residents were entered into the database.

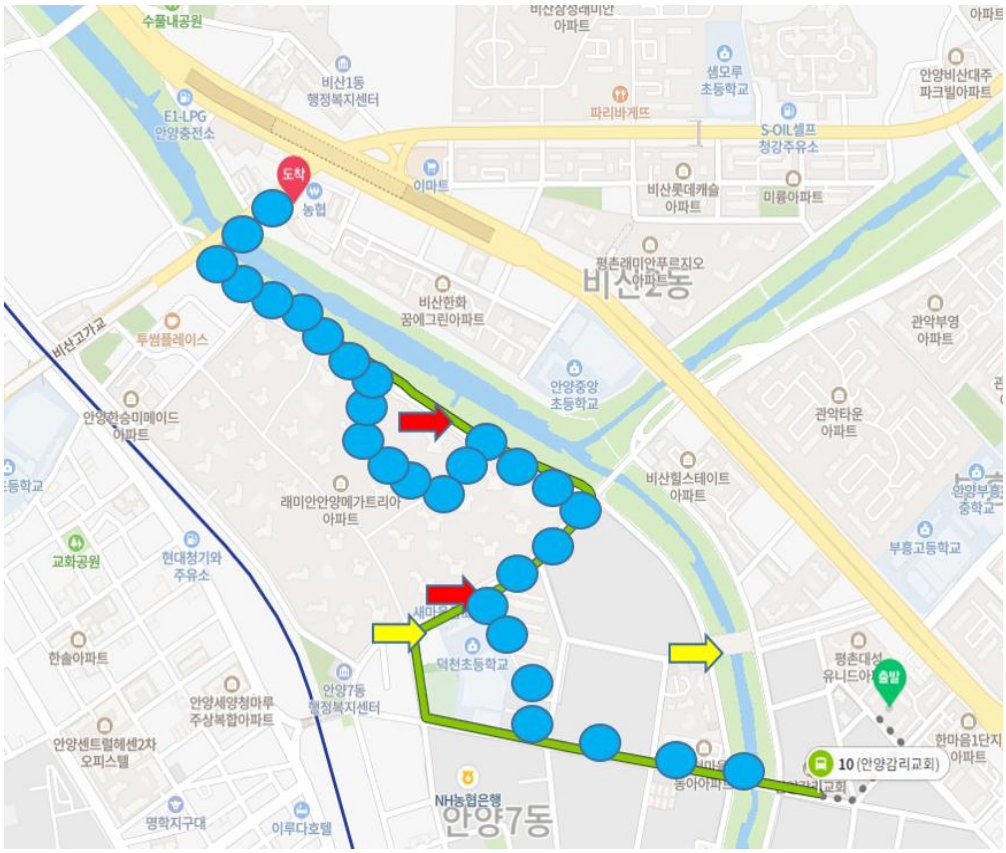

Fig. 3. Example of movement simulation

After verifying the consistency of this information and confirming that it is suggested as the actual application user's movement information, the simulation test was started. Set up a virtual path for the simulation experiment. In this process, two exception routes were entered between the starting point and the destination point. Information such as recent traffic conditions and building construction were input, and roads not used by users were additionally checked. More accurate technique was tried to select a bus stop or subway station near the destination, and searched for a route with two or more detours. For the proper setting of the experimental route, four points were set, with two positions each corresponding to the values of function 2 and function 3. The yellow and red markings in Figure 4 correspond to these. When moving the simulation, the information provided on the navigation route corresponds to the green solid line, and

the recommended information of the application proposed in the study corresponds to the continuation of the blue circle. If you check this line, you can confirm that the evacuation for the exceptional situation and the presentation as a corrected route were properly performed. According to the actual data, there are a total of 12 moves. During this time, a total of 524 guide information was analyzed, of which 12 times were duplicated data, and the remaining 512 times of data were analyzed. Among these figures, there were 19 factors corresponding to function 2 of the pathway and 26 factors corresponding to function 3 of the pathway. When the experiment was performed without applying for function 2 and function 3, the total time increased by 14%. In terms of accuracy, it can be seen that the case most similar to the movement information of actual local residents was brought.

IV. CONCLUSION AND DISCUSSION

In this study, we discussed how to reduce various exceptions on a regular basis and improve accuracy in application information guidance of location-based services. The process of analyzing the route movement of users is based on the premise that they travel the shortest route that local residents are familiar with. This study presented a method to collect data of users' various exceptions in real time and move them through a more stable system. There are two limitations of this study. The first is that there should be enough movement data that serves to create the underlying data. The second limitation is that the accuracy can be increased only when the number of exception cases reaches a certain number or more. It depends on the characteristics of the system used to refine the results and exception values based on data. Therefore, even if the data analyzed in the system is purified and processed efficiently, the reliability of the system will be lowered if the information is not accurate. In the simulation experiment of this study, analysis was performed on approximately 500 pieces of information, and it was confirmed that information can be used as common movement data when more than 80 pieces of information are collected from each movement route. Assuming that the simulation experiment is the movement of a refined route, the process replaced by the actual walking case is used. In this process, empirically, it can be seen that a value of 20% or more can provide a value corrected for accuracy and exception only when more than 110 movement data are collected. As a supplement to this point, it will be necessary to use the information of a specific person designated through the system's compensation system or information of a person who is presumed to be a local resident. In this method, it can be an alternative to introduce an algorithm that specifies a default value even if the number of data is small for information with high reliability. It is expected that a more accurate LBS system will emerge by handling various exceptions that occur in movement information and suggesting complementary routes.

ACKNOWLEDGEMENTS

This work was supported by the research fund of Hanyang Cyber University (2019)

REFERENCES

[1] https://terms.tta.or.kr/dictionary/dictionaryView.do?word_seq=047825-2.J

[2] Korea Internet & Security Agency, “Report on domestic and overseas LBS industry trends”, 2018.Software policy reaserch institute, Software-oriented society, 7,2021

[3] Tiwari, Shivendra, et al. "A survey on LBS: system architecture, trends and broad research areas." International Workshop on Databases in Networked Information Systems. Springer, Heidelberg, 2011.

[4] https://www.iamrenew.com/green-transportation/charge-electric-cars-google-maps/

[5] Geospatial Information Research Institute, Korea Land Information Corporation, “Study on ways to revitalize traditional markets using LBS”, 2016.

[6] Kushwaha, Amit, and Vineet Kushwaha. "Location based services using android mobile operating system." International Journal of Advances in Engineering & Technology 1.1 14.2011.

[7] Lee, W. J., Yoo, J. H., Lee, K. H., & Jang, E. G.. A Design and Implementation of Product Information Guide System Based on Beacon. Journal of the Korea Society of Computer and Information, 25(12), 119-126. 2020.

[8] Changbae Mun, & Hyunsuk Park. "A service for providing information related to local stores based on big data." Journal of the Korean Contents Association, Vol. 20, No. 2, pp. 561-571, 2020.

[9] Hyeongjun Kwon, & Kwangseok Hong. “Personalization of LBS using collaborative filtering-based recommendation system”. Journal of Internet Information Society, Vol. 11, No. 6, , 2020.

[10] Tilson, David, Kalle Lyytinen, and Ryan Baxter. "A framework for selecting a location based service (LBS) strategy and service portfolio." 37th Annual Hawaii International Conference on System Sciences, 2004. Proceedings of the. IEEE, 2004.

[11] Ververidis, Christopher, and G. Polyzos. "Mobile marketing using a location based service." Proceedings of the First International Conference on Mobile Business, Athens, Greece. 2002.

[12] Shao, Jun, Rongxing Lu, and Xiaodong Lin. "FINE: A fine-grained privacy-preserving location-based service framework for mobile devices." IEEE INFOCOM 2014-IEEE

[13] Chang-Bae Mun; Park Hyun-seok. Design of intelligent LBS service using big data distributed processing system. Journal of the Korean Contents Association, Vol. 19.2: pp.159-169.2019.

[14] Chang-Bae Mun; Park Hyun-seok. Sejun Yoon, A technique for improving the accuracy of LBS systems using big data. Proceedings of the Fall Conference of the Korean Society of Industrial-Academic Technology, pp343-346.2020.

[15] Changbae Mun, " Improving the Accuracy of Mobile LBS Applications." Advances in Dynamical Systems and Applications, Volume 16, Number 2, (2021) pp. 1145-1149

[16] Changbae Mun, " Improving the Accuracy of Mobile LBS Applications." International Journal of Engineering Research and Technology, Volume 13, Number 12, (2021) pp. 4151-4155