International Journal of Engineering & Technology, 7 (4.1) (2018) 113-117

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET

Research paperer

Development of an Asset Management System Integrated with GIS and K-Means Algorithm for Large Industrial Area

Ibid Athoillah[1*], Firda Pratiwi[2]

[1]Faculty of Computer Science, University of Dian Nuswantoro, Semarang, Indonesia
[2]PT Krakatau Steel (Persero) Tbk, Banten, Indonesia
**Corresponding author E-mail: ibidathoillah@gmail.com*

Abstract

The large industrial area has more than a thousand assets that make asset management activities to become more challenging. However, the current asset management in several industrial companies still records the asset information in paper-based documents. Therefore, the technical problems found in current practices in the field are the difficulties in determining the actual position of assets and which group of asset belongs accurately. To overcome those problems and to improve the asset management activities, an asset management system using GIS (Geographic Information System) and K-Means Algorithm is proposed. GIS to visualize the asset coordinate and condition on a geographical map, manage and store the assets detail information and K-Means Clustering Algorithm to cluster the assets based on their coordinate and simplifies the scattered asset. The system is developed using the Waterfall model. The system is native applications built using electron framework and can be accessed within industrial area network only, to keep the user convenience and security of asset data. The Leaflet used as interactive map software frameworks to visualize the industrial map, manage the geospatial data and mapped the industrial asset, including machinery, equipment, land, and building. For the initial operation, the system contains more than 700 assets data with 6 categories and 4 clusters. This system improves the assets accuracy, reduces the time of asset management practices, and the asset clusters can be used as decision support for industrial asset management activities.

Keywords: *Asset management;, Integrated system; GIS, K-means.*

1. Introduction

The purpose of asset management is to organize assets which accordance with companies' business needs and to provide support services so that they can operate it effectively [1]. This activity is very important since 85% of the total project expenses are spent on operation and maintenance (O&M) [2]. The current practices of asset management are paper-based comprise of manual assessment and turn out to be tedious, time-consuming, and has a tendency human error happen [2]. The best practices are not sufficiently recorded or documented [3], it requires an asset management system to support maintenance, acquisition, asset planning, and logistic [1]. Nowadays, the development of science and information technology is rapidly changing; the implementation of technology could improve asset management activity. Through asset management system integrated with the geographic information system, it can archive substantial benefit such as assets accurate, efficient and effective [4].

In the industrial area of this research, they have more than a thousand assets placed. The assets consist of building, machine, tools and equipment used to support their business. The entire asset data has been documented in the spreadsheet includes the assets status, coordinates, and the other attributes. However, with large asset data listed and old paper-based practices, it makes asset management activities ineffective and lack of visualization. The worker finds the asset hard just by looking at the coordinates only in the paper. Moreover, the assets document given is unsorted or random, the worker will back and forth when he wants to maintain a group of assets or specific asset. It will be troublesome and time-consuming.

To overcome those problems, the effective technique which optimizes asset management is required. The proposed technique to be implemented is using Geographic Information System (GIS) and clustering. Managing assets by mapping them provide a possibility monitoring the asset accurately [4]. However, the large scale of distributed assets and diverse environmental conditions is a challenge to asset inspection [5]. The clustering is useful for assets checking on a regular basis, and it used to plot the area or grouping the asset into many of plot. The assets cluster could lead to effectiveness and flexibility on asset monitoring activity. It enables teamwork to do asset monitoring simultaneously which division of labor can be divided according to the cluster, or gradually which the worker does not have to check in one day. This clustering assets technique using similar reliability distributions of an asset is enables addressing the maintenance optimization problem of all the assets belonging to the same cluster [6].

A large amount of asset data in GIS is required to minimize the time complexity of the clustering algorithm [7]. The proposed clustering technique used is K-Means Algorithm. The most widely used clustering techniques integrated with GIS is K-Means [7]. K-means [8] is well known for its efficiency in clustering large data sets, with little or no background knowledge. In practice, K-Means can be used as middleware to process smaller and specific pieces of data from huge and massive large sets. The implementation K-Means on asset management can simplify the large assets data scattered into a group of assets; those assets are grouped according to the closest distance. In this proposed system, the attribute used for clustering is coordinate and the asset which the closest coordinates to the predefined centroid become to one group or cluster.

Copyright © 2018 Authors. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The process of design and development system followed the Waterfall System Development Life Cycle. The Waterfall model is a linear and sequential process where each stage must be completed first before proceeding to the next stage [9]. The stage of waterfall consists of gathering the system requirement, design, implementation and unit testing, integration and system testing, and then the final stage is system operation and maintenance.

In this research, we proposed an Asset Management System Using GIS and K-Means Algorithm that can be used to organize the assets with map visualization get an accurate location of assets and makes asset cluster. This system uses well data integrity; it provided detail information such as location, description, condition, image, etc. That information can be used to supporting decision making in asset management activities.

2. Literature review

2.1. Geographic information system

A GIS is defined as "an information system that is designed to work with data referenced by spatial or geographic coordinates" [10]. The Geographic Information System (GIS) has been implemented to support analysis on asset management in the various fields. In the power sector, at Sultan Qaboos University (SQU), Oman [11], the research is integrated GIS with a smart grid. They used GIS to display electrical network and all its assets in real-time on a geographical map. The result of their development is a platform which provides the electricity utilities with a tool for monitoring, control, asset management, and demand-side management in a smart grid. In the transportation sector, Cristina Torres-Machi et al developed the transportation asset management system that supported with GIS. On the other hand, in Municipal Government, Maha Reda Al-Kasisbeh and Osama Abudayyeh [12] used GIS for managing public assets which included land, buildings, equipment, water, and sewage treatment plants, parks, roads, utilities, and bridges. The result of their study is a GIS providing automated techniques for managing asset information which increases data quality, integrity, and timeliness. From related study described above, we can conclude that GIS is effective to support decision making of asset management activity.

2.2. K-Means algorithm

K-Means is a simple partitioning clustering algorithm, it using iterative refinement technique to grouping the similar objects in the same cluster [13]. This algorithm belongs to unsupervised learning which receives input in the form of data without a class label. The aims of K-Means clustering is that minimizing the Sum of Squared Error (SSE) objective function to seek the optimal partition of the data [7]. There are various distance matric can be used to find the similarity of the object. According to Archana Singh et al, Euclidean distance metric gives the best result comparing to Minkowski and Manhattan distance [14]. Hence, this research used Euclidean distance.

Algorithmic Step for K-Means Clustering [15]:

Let X = {x1,x2,x3,……..,xn} be the set of data points and V = {v1,v2,…….,vc} be the set of centers. Allocate data into clusters randomly.

1. Select randomly 'c' center of the cluster (centroid).
2. Calculate the distance between each data point with the centroid using Euclidean distance as follows:

$$Dist_{XY} = \sqrt{\sum_{k=1}^{m}(X_{ij} - X_{gj})^2} \quad (1)$$

3. Allocate each data point to the nearest centroid.
4. Recalculate the new cluster center using:

$$v_i = (1/c_i)\sum_{j=1}^{c_i} x_i \quad (2)$$

Where, 'c_i' represent the number of data points in the ith cluster.

5. Recalculate the distance between each data point using new obtained cluster centers.
6. Repeat from steps 3 until no displacement data point.

3. Implementation

3.1. System requirement

The system requirement focuses on defining the functionalities, the limitation problem, and the purpose of the system. The requirements are in accordance with the agreement between the industry and the software developer. From the discussion, the first problem explained is the current assets management activity in the industry is lack of visualization, makes assets document is not valid enough especially coordinate of assets. The second problem is assets in the industrial area is scattered and need to be clustered, Asset cluster helps the monitoring activities can proceed gradually. Based on those problems and data gathered from discussion with the stakeholder, asset management system using GIS and the K-Means Clustering is proposed. This system provides the spatial data embedded with assets information on the offline geographical map. The spatial data and assets information can be managed with the user interface and imported from the asset document given.

3.2. System and software design

The system and software design phase determine the system requirements either hardware or software by forming the architecture of the overall system. The software design involves identifying and drawing abstractions basic software systems and relationships. The design relationship of actor interaction with the proposed system represents in use case diagram as shown in Figure 1.

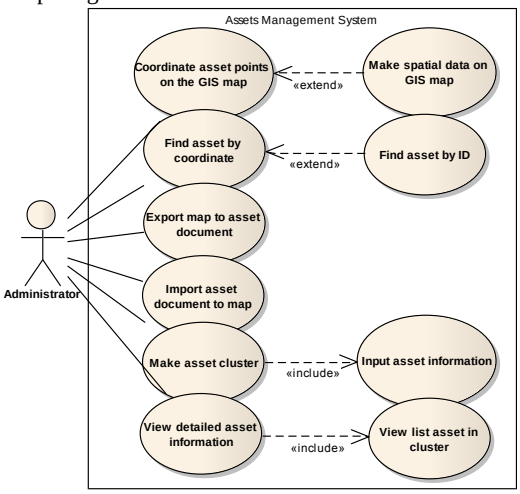

Fig. 1: Proposed system use case diagram

Figure 1 shows the system used by an actor named administrator. The administrator has the responsibility to manage the assets data and have access to all features in the system. The authentication is needed to get administrator access. The Proposed System Architecture was developed by using MySQL Database, ExpressJS Server, LeafletJS Map, and Electron (software framework). MySQL Database stores assets information and spatial data related in geojson format. ExpressJS Server is used as the bridge to connect the Desktop Apps with MySQL Database. LeafletJS Map is an open-source JavaScript library that is used to visualize the offline map of the industry. The LeafletJS can support the feature to build Geographic Information System (GIS) with an interactive map. The software framework Electron is to build a cross-platform desktop application. The detail of system architecture explained in Figure 2.

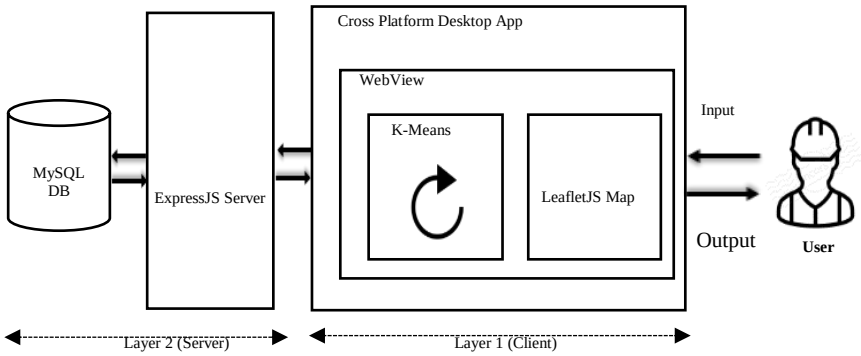

Fig. 2: Proposed system architecture

Figure 2 shows Layer 1 is desktop application interface for the user to manage the asset and spatial data. All process inside this layer run on the computer client. The process consists of managing spatial data related using map interface and the assets clustering process using K-Means also proceed in the client side. In Layer 2, the process consists of querying assets and spatial data information in the database. The request query from the client side is proceeding by the server and the result is returning to the client-side.

3.3. Implementation and unit testing

At this implementation and unit testing phase, Implementation of the design above is performed. This activity includes coding, data definition, building database, collecting the necessary data and files, and the other required activity. Along the implementation, unit testing is necessary to detect software error and failures, so the defects may be discovered and corrected. The testing technique used is white-box and black-box testing. White Box Testing [16] is a testing technique based on the internal structure of the component or system. In the implementation, the programmer does the white box testing. It is necessary for a tester to have good programming knowledge and have better understand the source code. The Black Box Testing [17] is carried out by testers who don't have knowledge of the internal structure and have no idea about the source code of the system. The tester only knows how to operate the system. This technique is used to validate that the functional requirement is provided correct output.

3.3.1. K-means clustering implementation

a) Data collection

We collected the data from the HSE division in the related industry. The position asset on the map gathered using Garmin. The data is in the form of a table that has a constituent component as follows:

1. Has 4 attributes namely asset id, description, and coordinate latitude and longitude.
2. The number of sample instances is 15 assets to clustering, the list of data as shown in Table 1.

Table 1: Sample asset data for the clustering process

ID	Description	Lat	Long
1	Loading Arm Type: E110-212	-6.xxx417	110.xxx039
2	Ac Split Merk: Daikin 3pk	-6.xxx963	110.xxx632
3	Air Compressor M. Ingersoll Rand M234.	-6.xxx144	110.xxx062
4	Air Compressor M. Ingersoll Rand M234.	-6.xxx374	110.xxx088
5	Automatic Tank Gauging	-6.xxx262	110.xxx861
6	Bottom Loader Dia 4"	-6.xxx008	110.xxx227
7	Bottom Loader Dia 4"	-6.xxx357	110.xxx948
8	Bulk Meter	-6.xxx564	110.xxx262
9	Electric Actuator	-6.xxx650	110.xxx377
10	Fire Alarm Box	-6.xxx539	110.xxx533
11	Flexible Hoses 3" 60ft	-6.xxx645	110.xxx753
12	Flexible Hoses 3" 60ft	-6.xxx417	110.xxx433
13	Floating Hose 16"X30"	-6.xxx160	110.xxx632
14	Floating Hose 16"X35" 1"	-6.xxx795	110.xxx704
15	Floating Hose 16"X35" 2"	-6.xxx901	110.xxx495

b) Data modeling

When the format position gathered is in Degrees Minutes Seconds. The DMS format should be converted into numeric value or latitude and longitude to do a calculation using K-Means. Because of the monitoring executed on the square map, the assets are divided into 4 clusters. The centroid is predefined as desired by the monitoring implementer. Table 2 shows the detailed coordinate of each centroid and Figure 3 shows the position of centroid selected in the map.

Table 2: Coordinate each centroid

Centroid	Latitude	Longitude
Centroid 1	-6.xxx548846	110.xxx653
Centroid 2	-6.xxx521141	110.xxx375
Centroid 3	-6.xxx832144	110.xxx393
Centroid 4	-6.xxx246407	110.xxx389

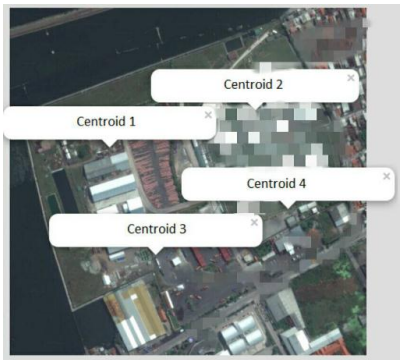

Fig. 3: Coordinate centroid on map

c) Clustering process

By using Euclidean distance formula, Table 3 shows the distance between each data with the centroid. The shaded column shows the nearest centroid with the data.

Table 3: 1st Iteration distance between data with centroid

ID	Centroid 1	Centroid 2	Centroid 3	Centroid 4
1	0.000408	0.001609	0.001459	0.002492
2	0.002064	0.000512	0.002243	0.001490
3	0.002442	0.000927	0.002373	0.001150
4	0.002571	0.001986	0.001756	0.000326
5	0.003286	0.002289	0.002533	0.000472
6	0.001568	0.002739	0.000242	0.002292
7	0.000352	0.001654	0.001541	0.002598
8	0.001887	0.000120	0.002429	0.002025
9	0.000731	0.001507	0.001182	0.002098
10	0.001325	0.002187	0.000324	0.001879
11	0.001104	0.001284	0.001241	0.001743
12	0.000256	0.002139	0.001710	0.003070
13	0.002017	0.000689	0.002081	0.001324
14	0.001630	0.002371	0.000313	0.001771
15	0.001593	0.002538	0.000123	0.002004

The next iteration is calculating Euclidian distance using new centroid that gathered from the average of each cluster. This calculation finishes at the second iteration because there was no displacement found.

d) Classified result cluster

From Table 3, we can classify the asset by comparing the distance of each data to the centroid. Table 4 shows the nearest asset data to the centroid after compared.

Table 4: Classified result cluster

Centroid	Nearest asset data
1	1, 7, 9,11, 12
2	2,3,8,13
3	6,10,14,15
4	4,5

e) Validating k-means clustering of proposed system

To validate the K-Means Algorithm that implemented in the proposed system, we use the above result cluster as a comparison. With the same data collection, Figure 4 shows the clustering process executed on the proposed system.

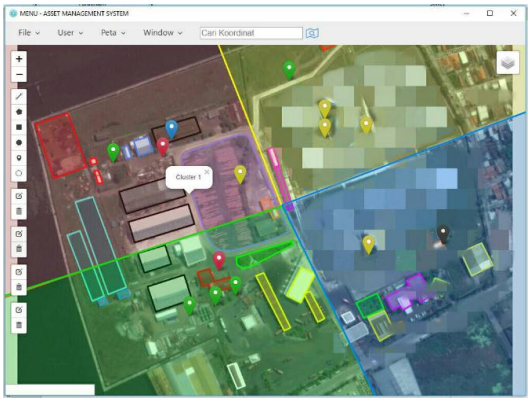

Fig. 4: Proposed system clustering result

The system clustering result:

1. Cluster 1 (Red) with asset id 1,7,9,11,12
2. Cluster 2 (Yellow) with asset id 2,3,8,13
3. Cluster 3 (Green) with asset id 6,10,14,15
4. Cluster 4 (Blue) with asset id 4,5

From those result, we can conclude that implementation K-Means in this proposed system is correctly implemented because the cluster is same with the result in Table 4.

3.4. Integration and system testing

After the system design architecture and all requirement specification are implemented, the next phase is integrating the data assets to the proposed system. The integration is using the real data from the related industry. The user tested the system testing; begin from installing the system to the user computer until operating each function of the system. This stage is required to check whether the software system is made according to the customer needs.

3.5. Operation and maintenance

After the system integrated and tested, the final stage is Operation and Maintenance. The system is installed on each computer in the office of the industry and ready to operate into work activities. The maintenance will carry out when improvements and feature additions are required, or when several functions that need to repair.

4. Result application

The system design and architecture above have been implemented. For the initial operation, more than 700 asset data have been included automatically using the import document to map functionalities. The assets imported consist of land, building, machine, tools, equipment, etc. As shown in Figure 5, Marker represents the location of the assets and polygon, rectangle, and the circle represents specific spot and building assets. Each asset marker on the map has a specific color depending on the condition of the assets. Table 5 shows six unique asset colors following by the condition code and their description. In this system, the user can find the assets by its coordinate and the asset identity; the map will redirect and focus on the asset desired. The detailed information of the asset also can be known by clicking on the icon on the map. The Filter of cluster functionalities can use by selecting the layer that wants to be filtered and choosing one of list cluster desired. Then the system will show the list of assets that only belongs to the selected cluster. That list of assets can use as decision making in asset inspection. The implemented system helps to perform better asset management. By the increase of asset accuracy, we can know where the exact location of the asset. The system supports the asset inspection in the decision of planned routes. With many assets, the inspection activities should be done simultaneously, and by use of the system, the work can be divided according to the cluster created. Easy to maintain the asset, how the current condition of the asset whether the asset is feasible or dangerous to use. The asset conditions greatly affect the workers' safety. Therefore, before workers use assets, they must know that assets are feasible or dangerous to use, that through this system, they can access all assets.

Table 5: Marker color description

No	Marker color	Condition Code	Description
1		FC	Reserved
2		FR	Damaged Condition
3		DB	Good Condition
4		TR	Bad Condition
5		DS	Moderate Condition
6		FX	Not found

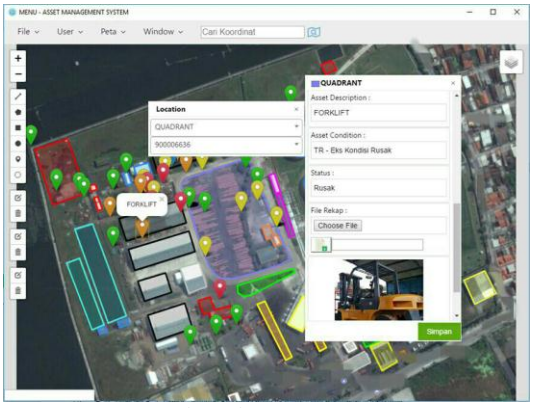

Fig. 5: Interface of the implemented system

5. Conclusion

The development of the system by using the waterfall model has performed. The asset management system for the large industrial area is well developed. The final product is in the form of a cross-platform desktop application. Although the Leaflet GIS is the library for the web-based application, this development makes the use of the electron framework to build the desktop apps which can run on cross-platform. The integration of Leaflet GIS Library to the system has performed, and the data is well integrated. The system usage is private software that can run offline without internet connection. However, the installation is required the MySQL connection. To synchronize the data between computers, it can use the same connection of database and LAN (Local Area Network). The asset management system proven can support asset inspection, managing the assets information with visualization and paperless. Asset clustering methods are done by grouping the data into four clusters based on their coordinate, and finally, the clustered assets displayed correctly in the form of maps with markers that can use for asset inspection decision making. This kind of technology is needed to be implemented in any industrial sector.

References

[1] Hastings NAJ, *Physical Asset Management: With an Introduction to ISO55000*, 2nd ed. Springer Berlin Heidelberg, (2015).

[2] Aslam HM & Tarmizi HA, "Intelligent BIM record model for effective asset management of constructed facility," *MATEC Web Conf.*, Vol.195, (2018), p.7.

[3] Emmanouilidis C & Komonen K, "Physical Asset Management Practices in Industry: Comparisons between Greece and Other EU Countries BT-Advances in Production Management Systems. Sustainable Production and Service Supply Chains", *Conference on Advances in Production Management Systems*, (2013), pp.509-516.

[4] Hari Ginardi RV, Gunawan W & Wardana SR, "WebGIS for Asset Management of Land and Building of Madiun City Government", *Procedia Computer Science*, Vol.124, (2017), pp.437-443.

[5] Jin C *et al.*, "A comprehensive framework of factory-to-factory dynamic fleet-level prognostics and operation management for geographically distributed assets", 2015 *IEEE International Conference on Automation Science and Engineering*, (2015).

[6] Cannarile F, Compare M, Di Maio F & Zio E, "Handling reliability big data: A similarity-based approach for clustering a large fleet of assets", *Saf. Reliab. Complex Eng. Syst.*, (2015).

[7] Hamfelt A, Karlsson M, Thierfelder T & Valkovsky V, *Beyond K-means: Clusters identification for GIS*, Lecture Notes in Geoinformation and Cartography, (2011).

[8] Arora P, Deepali, and Varshney S, "Analysis of K-Means and K-Medoids Algorithm for Big Data," *Physics Procedia*, Vol.78, (2016), pp.507-512.

[9] DESPA ML, "Comparative study on software development methodologies", *Database Syst. J.*, Vol.V, No.3, (2014), pp.37-56.

[10] Chrisman NR, "Review Paper What Does ` GIS ' Mean ?", *Main*, Vol.3, No.2, (1999), pp.175–186.

[11] Ashkezari AD, Hosseinzadeh N, Chebli A & Albadi M, "Development of an enterprise Geographic Information System (GIS) integrated with smart grid", *Sustain. Energy, Grids Networks*, Vol.14, (2018), pp.25-34.

[12] Al-Kasisbeh MR & Abudayyeh O, "Municipality asset management: Asset types and effective management decision using GIS", *Construction Research Congress 2018: Infrastructure and Facility Management-Selected Papers from the Construction Research Congress 2018*, (2018).

[13] Aggarwal CC & Reddy CK, *Data Clustering: Algorithms and Applications*, CRC Press, (2018).

[14] Singh A, Yadav A & Rana A, "K-means with Three different Distance Metrics", *Int. J. Comput. Appl.*, Vol.67, No.10, (2013), pp.13-17.

[15] Rajaguru H & Prabhakar SK, *KNN Classifier and K-Means Clustering for Robust Classification of Epilepsy from EEG Signals. A Detailed Analysis*, Anchor Academic Publishing, (2017).

[16] Kumar M, Singh SK & Dwivedi R, "A Comparative Study of Black Box Testing and White Box Testing Techniques", *Int. J. Adv. Res. Comput. Sci. Manag. Stud.*, Vol.3, No.6, (2015), pp.12-15.

[17] Jan SR, Shah STU, Johar ZU, Shah Y & Khan F, "An Innovative Approach to Investigate Various Software Testing Techniques and Strategies", *Int. J. Sci. Res. Sci. Eng. Technol.*, Vol.2, No.2, (2016), pp.682-689.