

# An Efficient Algorithm To Detect The Nearest Location Of A Map For A Given Theme

M. A. P. Chamikara, Y. P. R. D. Yapa, S. R. Kodituwakku, J. Gunathilake

**Abstract:** -This paper presents an algorithm for efficient detection of the nearest police station for a specific position(coordinates). The algorithm is based on Geographic Information System (GIS), Geographic Positioning Systems (GPS) and the J48 classification algorithm. In other words, the proposed algorithm is an effective integration of these three. The salient feature of this algorithm is that it facilitates the detection of the current ground position on a digital map accurately and easily. The algorithm is tested by detecting the nearest police station of a given coordinate of the Sri Lankan map. The test results reported an average accuracy of 87.92%.

**Index Terms:** - Artificial Intelligence, Classification Algorithms, Data Mining, Geographic Information Systems (GIS), Intelligent Policing,

## 1 INTRODUCTION

Sri Lanka Police currently uses a manual mapping system to locate and analyse criminal incidents. In this procedure, the investigation of a crime is aided by the analysis of a hand-drawn map. However, the accuracy and effectiveness of crime analysis can be improved significantly with the use of the current advancements in information technology, enabling police to arrive at credible decisions within a comparably short span of time. SL-SecureNet [7] is a GIS based Crime Analysis System, which aims to implement most of the crime analysis tasks in an automated manner with a higher degree of accuracy. It is an ongoing research at the Post Graduate Institute of Science, University of Peradeniya, Sri Lanka. The project is granted by the National Research Council, Sri Lanka. A geographic information system (GIS) integrates hardware, software, and data for capturing, managing, analysing and displaying all forms of geographically referenced information. GIS allows the data to be viewed, understood, questioned, interpreted and visualized in numerous ways revealing relationships, patterns, and trends in the form of maps, globes, reports and charts[8]. A Global Positioning System (GPS) device communicates with four or more satellites to determine its exact location coordinates (latitude and longitude) anywhere on the Earth. It works in any weather as long as the device has a clear line of sight to the satellites. A-GPS, Assisted Global Positioning System, works on the same principles as the GPS.

However, it obtains coordinates data via a network system called assistant servers, unlike the GPS where coordinate data is obtained directly from satellites. Upon request, assistant servers fetch GIS data from satellites and feed them to the user agents, such as computers or mobile devices, which eventually consume the data [9]. SL-SecureNet employs many data mining techniques such as spatial analysis, periodic pattern analysis, time series analysis, cluster analysis, classification, artificial intelligence, and tools such as Crime locator, Crime clock, Periodic pattern visualiser, Public crime map, Hotspot detection, Nearest police station detection, Crime comparator, Outbreaks detection and so forth [7]. These functionalities help the police personals to arrange the crime records and analyse them in an easy manner. Nearest police station detection (NPSD) is one of the very important tools from the few mentioned above. NPSD tool provides the capability of identifying the nearest police station to a particular location in a user friendly and an instant manner. NPSD plays a vital role in prevention and investigation of crimes alike that it enables to locate the nearest police station to a particular location. This proves to be critical in situations which need immediate police assistance. For example, in case of a road accident, it can help the victims or witnesses to find out the nearest police station in order to call for assistance. Similarly the police might need to call for immediate backup from the nearest police station, when a suspect is on pursuit. NPSD can be an ideal tool for such situations. However, implementing this system in Sri Lanka is subject to several challenges. Unavailability of GIS information, mainly the road map layers of some rural areas, makes relying solely on GIS data not reliable for the process. NPSD tool comes with the longitudinal and latitudinal data pre-configured in to it. Due to this, it does not have to retrieve GIS information, although, it has to be trained for longitudinal and latitudinal information previously. This paper discusses an efficient method of integrating the J48 classification algorithm into a GIS system to address the above mentioned task. Our test results have shown that, NPSD has an average accuracy of 87.92% in identifying the nearest police station to a particular location.

- **M.A.P. Chamikara**, Post Graduate Institute of Science (PGIS), Faculty of Science, University of Peradeniya, Peradeniya, Sri Lanka, PH-+94718016715, (e-mail: [pathumchamikara@gmail.com](mailto:pathumchamikara@gmail.com)).
- **Y.P.R.D. Yapa**, Department of Statistics and Computer Science, Faculty of Science, University of Peradeniya, Peradeniya, Sri Lanka, PH-+94 81 239 (4642), (e-mail: [roshany@pdn.ac.lk](mailto:roshany@pdn.ac.lk)).
- **S.R. Kodituwakku**, Department of Statistics and Computer Science, Faculty of Science, University of Peradeniya, Peradeniya, Sri Lanka, PH-+94 81 239 (4641), (e-mail: [salukak@pdn.ac.lk](mailto:salukak@pdn.ac.lk)).
- **J. Gunathilake**, Department of Geology, Faculty of Science, University of Peradeniya, Peradeniya, Sri Lanka, PH-+94777313063, (e-mail: [aaikq@yahoo.com](mailto:aaikq@yahoo.com)).

## 2 Materials and methodology

### 2.1 Materials

SL-SecureNet system is composed of many Application Programming Interfaces (APIs), tools, frameworks, and libraries. This section includes a description of some of them which aids the integration of the NPSD tool into SL-SecureNet.

*i. J48 Decision Tree Algorithm*

J48 decision tree is a predictive machine-learning model that decides the target value (dependent variable) of a new sample based on various attribute values of the available data [1]. J48 is an implementation of the C4.5 algorithm under the WEKA data mining platform [6].

*ii. Openlayers*

Openlayers is a free and open source API developed using pure JavaScript for displaying map data in most modern web browsers without any server-side dependencies. Google Maps and MSN Virtual Earth APIs are some of the similar APIs with Openlayers [2].

*iii. Google Maps*

Google Maps is a map service provided by Google and that can be viewed in a web browser. It can be used free of charge for non-commercial purposes. Google Maps facilitates many services such as the Google Maps website, Google Ride Finder, Google Transit, and maps embedded on third-party websites via the Google Maps API. It also provides the capability of viewing local business information, including business locations, contact information, and driving directions depending on the location of a particular recipient [5].

*iv. Spring MVC.*

MVC is a framework abbreviated for Model-View-Controller which was made to develop the J2EE applications easier. Spring MVC shows how to move beyond Java Server Pages and uses other templating languages such as Velocity and FreeMarker [3]. Spring MVC is used as the URL routing mechanism for SL-SecureNet. The life cycle of a request in Spring MVC is depicted in Fig.1.

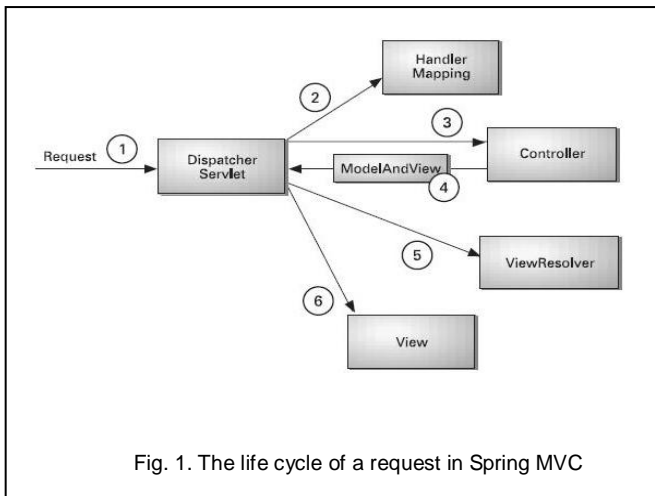


Fig. 1. The life cycle of a request in Spring MVC

*v. Weka*

Weka is a collection of machine learning algorithms for data mining tasks. It contains tools for data pre-processing, classification, regression, clustering, association rules, and visualization. These algorithms can either be applied directly from Weka itself, or can be called from Java source code [4].

*vi. Map Layers*

Spatial datasets are normally handled as layers in Geographic Information Systems. In order to help organization, geographical features are categorized under various themes such as roads, buildings, lakes, terrains and so on. Data under a particular theme is represented as an individual map layer [7]. In SL-SecureNet, Google maps serve as the base layer in referring to which the other features are located.

**2.2 Methodology**

NPSD tool has been implemented using a well ordered flow of events. Each of these events has been integrated to the NPSD tool with an important meaning. This section describes the flow of events in detail. Fig.2 depicts the architecture of the NPSD tool. A conventional web browser can be used to view the base Google Map rendered through the OpenLayers API. When the user requests a particular map to be rendered, the system communicates with the OpenLayers API, which in turn communicates with the Google Map service to render the requested map as the base layer.

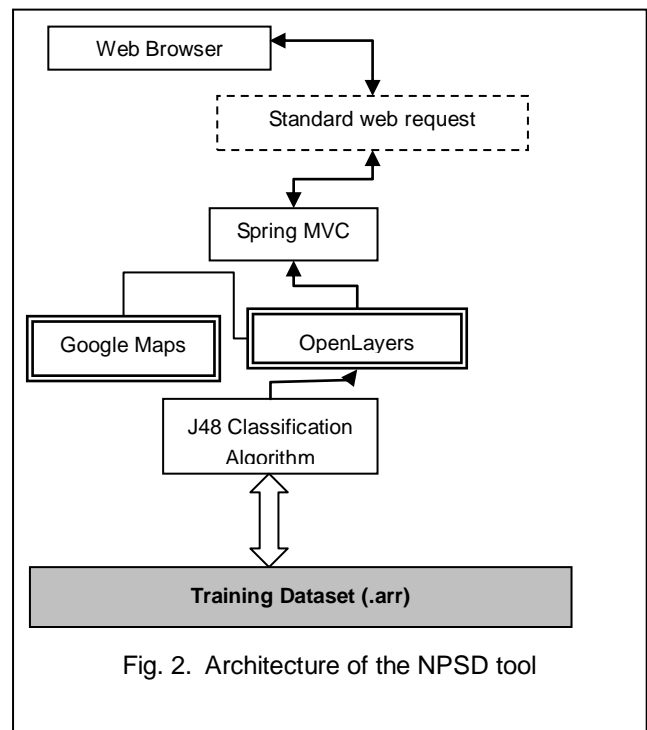


Fig. 2. Architecture of the NPSD tool

In the program, user action on the web browser triggers an HTTP request which is sent to Spring MVC. Users find the nearest police station to a location by clicking on the location of the map displayed on the web browser. If the user is using the mobile version, the current location will be tracked through A-GPS. Then the longitude and latitude information of that point get bound to an HTTP request as parameters. The HTTP request is sent to Spring MVC which filters the URL and checks for its authenticity. If it passes, it is routed to the logic controller, the J48 classification algorithm. J48 algorithm classifies the data (longitude latitude data) of the URL into the predefined group, i.e. the nearest police station according to the J48 pruned tree which is depicted in Figure 3. Then the servlet which corresponds to the NPSD algorithm appends this data to a HTTP POST request that is sent to the browser

through the Spring MVC URL routing mechanism. This request which is sent to the Spring MVC core is filtered and directed to the J48 classification algorithm, which is the main component of NPSD. The J48 algorithm has been pre-trained for considerable amount of instances per each area of interest on the map. These instances include the coordinates (longitudes and latitudes) and the nominal class values corresponding to the nearest police stations and nearest police stations are taken to be the class values. Thus, the training dataset contains a set of predefined police stations (classes) and each of the coordinates of the training dataset has been categorized under the relevant class. This will generate a classifier tree, which is to be used in the testing phase as shown in Figure 3.

```

lat <= 727699512670.169984
| lon <= 8066865921020.500992
| | lon <= 8063913345336.899584
| | | lat <= 724957941792.540032: HINDAGALA (9.0)
| | | lat > 724957941792.540032: PERADENIYA (21.0)
| | | lon > 8063913345336.899584: AMPITIYA (17.0/1.0)
| | lon > 8066865921020.500992
| | lat <= 726209549237.719936: THALATHUOYA (16.0)
| | lat > 726209549237.719936
| | | lat <= 727324897944.889984
| | | | lon <= 8068136215209.99936: ILUKMODARA (2.0)
| | | | lon > 8068136215209.99936: GURUDENIYA (15.0)
| | | lat > 727324897944.889984: ILUKMODARA (14.0/1.0)
lat > 727699512670.169984
| lon <= 8064291000366.198784
| | lon <= 8062025070190.399488: HALLOLUWA (10.0)
| | lon > 8062025070190.399488: KANDY (21.0/1.0)
| | lon > 8064291000366.198784
| | | lat <= 728891447798.729984
| | | | lon <= 8067586898803.699712: TENNEKUMBURA (14.0)
| | | | lon > 8067586898803.699712: KUNDASALA (12.0)
| | | lat > 728891447798.729984
| | | | lat <= 730747397909.869952: SIRIMALWATTA (13.0)
| | | | lat > 730747397909.869952: POLGOLLA (11.0)
    
```

Fig.3. J48 Pruned

J48 algorithm which is well trained in this way is allowed to access the Training dataset in response to the web request sent by the user. Once a user sends a web request for the nearest police station, the trained J48 classification algorithm maps the coordinates to the correct class. Then that class (police station), coordinate points is categorized under, is sent to the user as the response from the system. Figure 4 depicts a screenshot of the NPSD tool.

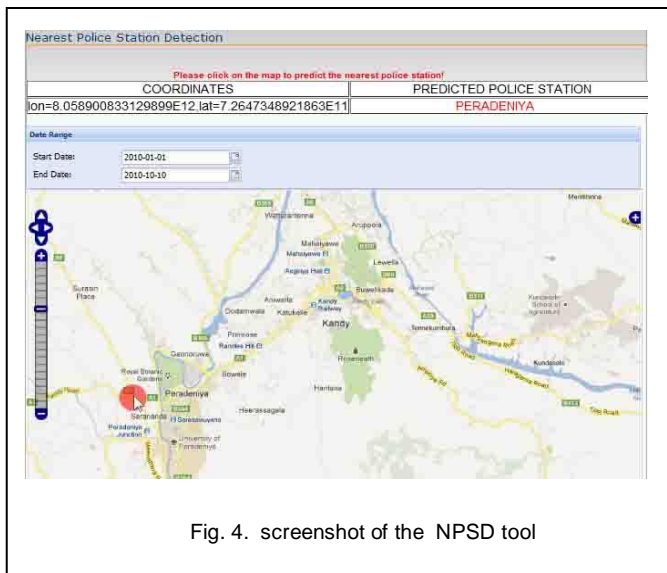


Fig. 4. screenshot of the NPSD tool

NPSD tool is a responsive web based tool where it has the capability of automatically adjusting the system to the mobile version whenever it is accessed via a mobile device by clicking the installed NPSD application or accessing it via a web page available in the mobile device. Figure 5 shows the home screen of the mobile version of the NPSD tool where the current location of a particular user is tracked through the A-GPS service provided by the service providers. It is important that the mobile device has the GPS capability or else the user will have to trigger the event sending the coordinate data by clicking on the map. To do that the user should have a transparent idea about his current location.

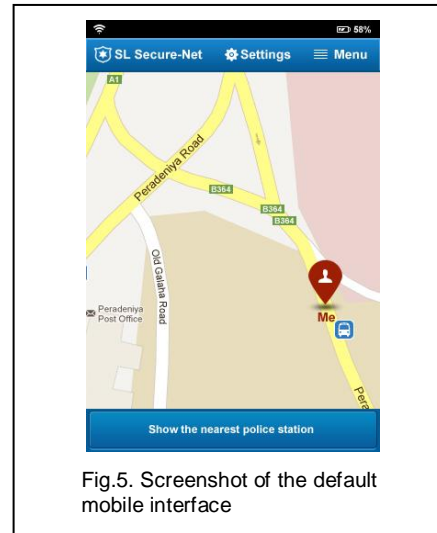


Fig.5. Screenshot of the default mobile interface

Current position of the user is indicated as the red icon annotated as 'Me'. This points to the relevant location of the user in the Google Map and is updated continuously as the user moves using A-GPS. User can make the request using the 'Show the nearest police station' button which triggers the process. As a result, the nearest police station is indicated on the map by another locator icon with name of the relevant police station. For the places where GIS information is available, the tool can indicate the path to the nearest police station as well. For other places, still the positions can be indicated on the map. In the mobile user interface, the user is presented with the ability to call the nearest police station using the 'Make a call' button. User can redo the search to detect the next nearest police station using the 'Show next' button. Otherwise, he/she can return to the previous result by using the 'Show previous' button shown in Fig. 6.

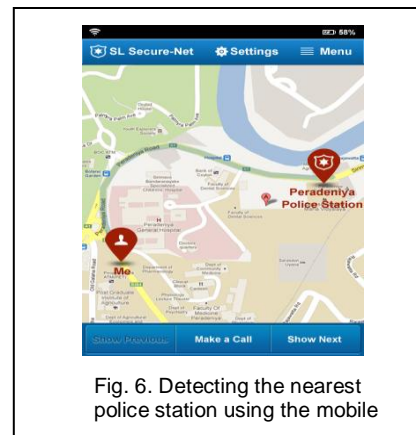


Fig. 6. Detecting the nearest police station using the mobile

### 3 Results and discussion

The area considered for testing is namely Kandy region which is a police region located in Sri Lanka. The dataset considered for the Kandy region consists of 175 instances. Each instance has a longitude, latitude and a nominal class value (a police station) in relation to that. Figure 7 shows the distribution of the instances over the class values where the dataset is composed of 12 classes (police stations) and the pruned tree obtained by training the J48 classification algorithm is depicted in Figure 3. The test results indicated that the system can be used to identify the nearest police station effectively. Here, the results of one area which the NPSD tool was tested for is attached.

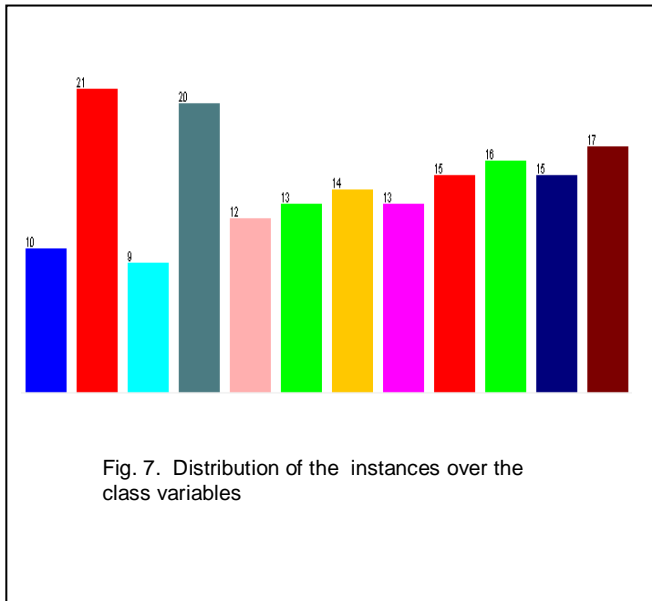


Fig. 7. Distribution of the instances over the class variables

The main concept behind the NSPD tool is the classification. The classification can be carried out anticipating four types of possibilities: true positive (TP), true negative (TN), false positive (FP) and false negative (FN). Those who test positive for a condition which is positive is called a TP. Those who test positive for a condition which is negative is called a FP. Those who test negative for a condition which is negative is called a TN. Those who test negative for a condition which is positive is called a FN [10]. However, NSPD tool would derive only two types of decisions from the above four: TP and FP. It is because of the fact that any coordinate point can always be classified to its corresponding class, i.e. the nearest police station. TP assigns a coordinate point to a correct nearest police station while FP assigns one to an incorrect nearest police station. When the percentage of the training data set is high, the data set includes a large portion of data from the main data sample and in effect the algorithm will be trained for a large portion of the dataset compared to the testing set. So, there is a high probability that the classifiers accuracy gets increased, and also when the percentage of the training data set is high, the dataset may contain more data from different classes. Therefore, the classifier will be trained to classify the dataset into classes easily because in the training phase it has been trained for a large diversity of instances. The performance plot of the J48 classification algorithm where the above sample dataset is used is depicted in Figure 8.

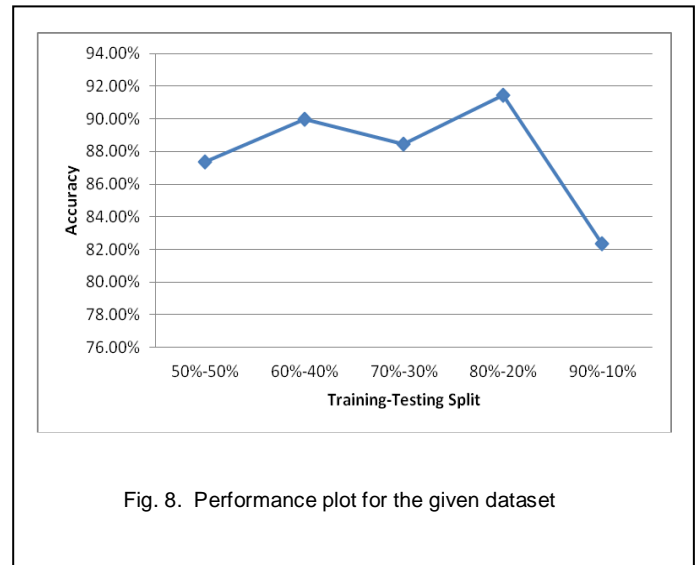


Fig. 8. Performance plot for the given dataset

### 4 Conclusion

The integration of J48 classification algorithm with GIS and A-GPS/GPS techniques has provided a successful method of identifying the nearest police stations effectively and is also applicable for places where the GIS data of the road map layer is not available. According to the performance plot shown in Fig. 8, the NPSD tool has showed an average accuracy of 87.92%. Although the algorithm is tested for detecting the nearest police station for a given coordinate position, it can be used to detect the nearest location of a map for a given theme as well.

### ACKNOWLEDGMENT

This work is funded by the National Research Council (NRC) of Sri Lanka [Grant number: 11-071]

### REFERENCES

- [1] (2010) Classification Methods. [Online]. Available : <http://www.d.umn.edu/~padhy005/Chapter5.html>
- [2] (2010) OpenLayers. [Online]. Available : <http://openlayers.org>.
- [3] Craig Walls & Ryan Breidenbach, Spring in Action, 2nd ed, Manning Publications, USA 2005.
- [4] (2011) Weka.[Online]. Available: <http://www.cs.waikato.ac.nz/ml/weka/>
- [5] (2012) Google Maps [Online] Available: <http://support.google.com/maps/bin/answer.py?hl=en&to pic=1687350&answer=144352>
- [6] S. Moore, D. D'Addario, J. Kurinskas, and G. M. Weiss. Are Decision Trees Always Greener on the Open (Source) Side of the Fence? In Proceedings of the 2009 International Conference on Data Mining, CSREA Press, 185-188, 2009.
- [7] M.A.P. Chamikara, Y.P.R.D. Yapa, S.R. Kodituwakku, J. Gunathilake, SL-SecureNet: INTELLIGENT POLICING USING DATA MINING TECHNIQUES, International



Journal of Soft Computing and Engineering (IJSCE),  
Volume-2, Issue-1, March 2012, p175-180

- [8] (2012) ESRI [Online] Available:  
<http://www.esri.com/what-is-gis/overview.html>
- [9] (2012) A-GPS Vs GPS [Online] Available:  
[http://www.diffen.com/difference/A-GPS\\_vs\\_GPS](http://www.diffen.com/difference/A-GPS_vs_GPS)
- [10] (2012) Medical Testing by Jerry Johnson [Online]  
Available:  
<http://wolfweb.unr.edu/homepage/jerryj/NNN/MedicalTesting.pdf>



**M. A. Pathum Chamikara** is working as a research assistant at the Post Graduate Institute of Science (PGIS), University of Peradeniya, Sri Lanka. He received his BSc (Special) degree in Computer Science, University of Peradeniya, Sri Lanka (2010). His research interests include Crime analysis, GIS (Geographic Information Systems), image processing, computer vision and artificial intelligence.



**Y.P. R. Darshana Yapa** is a senior lecturer at the Department of Statistics and Computer Science, University of Peradeniya, Sri Lanka. His research interests include image processing, computer vision and GIS (Geographical Information Systems)



**S. R. Kodituwakku** is a professor at the Department of Statistics and Computer Science, University of Peradeniya, Sri Lanka. His research interests include database systems, distributed computing, Role based Access Control systems, and software engineering.



**Jagath Gunathilake** is a senior lecturer at the Department of Geology, University of Peradeniya, Sri Lanka. His research interests include Natural resource management, GIS (Geographical Information Systems) and RS (Remote Sensing) systems.