ISSN: 2454-4248
719 – 721

Irregularity Finding in Roads Conditions using Data Mining: A Survey

Sunny Gandhi
Asst. Prof.
DMIETR, Wardha
sunnygandhi09@gmail.com

Anand Sharma
Asst. Prof.
SSGMCE, Shegaon
sharma.anand2008@gmail.com

`

Abstract- Road conditions play a vital role now days. Irregularity in road surface can cause accidents, vehicle failure and discomfort in drivers and passengers. Governments spend lots of amount every year in maintenance of roads for keeping roads in proper condition. But more maintenance work can increase the traffic, causing disturbance in road users. To avoid disturbances caused by road irregularity,this system can detect road irregularity using Smartphone sensors. The approach is based on data mining. In this, it used scikit-learn, a python module, and Weka, as tools for data-mining. All cleaning data process was made using python language. The final outputs show that it is possible to find out road irregularity.

*****

1. Introduction

Road conditions play a vital role now days. Road conditions play a vital role now days. Irregularity in road surface can cause accidents, vehicle failure and discomfort in drivers and passengers. Governments spend lots of amount every year in maintenance of roads for keeping roads in proper condition. But more maintenance work can increase the traffic, causing disturbance in road users.

A road irregularity tracking system can help in planning and scheduling road maintenance in a way that decreases the impact in people daily livewhich may decrease the number of accidents and costs in vehicles maintenance. Also, when planning a tour, information of roads condition can reduce the disturbance of passengers and drivers. [1, 2] However, analyzing road conditionrequires expensive and equipments and vehicles that cost more money, it also required trained workers to operate these machines [3]. Most drivers have Smartphone;have sensing features, namely location sensors [4]. Also, Smartphone are part of our daily routine, with most people owning, at least a Smartphone. Therefore these Smartphone could be used as road monitoring equipment.In this survey this system works on data collected by Smartphone sensors, automatically detects road irregularities using data mining.

The use of Smartphone devices imposes several problems and limitations when compared to the more expensive and specialized road-monitoring equipment. Specifically, Smartphone uses less costly sensors which reveal less sensitivity and readings when compared to upgraded equipment. This is correct for GPS sensors. On the other hand Smartphone and sensor makes and outputs as different sensors typically measure different values for the same physical action [3]. This difference is not only related to Smartphone and sensors.

The automotive department suffers from the same difficulty, with many different vehicle models, different suspension types and tire sizes. All this has direct impact on the collected data, thus existing solutions that detect and categorize road irregularities based on pre-defined threshold values are not accurate for real world conditions [5, 6, 7].

The GPS sensors of Smartphone typically have an inaccurate ranging from a couple to tens of meters [8, 9], meaning that the same location can have different latitude and longitude in different situations. Driving speed has a considerable role in vehicle conditions, creating higher vibration amplitude[10].

2. Related Work

When we discuss about road condition with monitoring systems, we generally talk about data centered systems. Data is typically collected by Smartphone and analyzed to understand the exact problem, and to obtain knowledge about it. However, raw data is difficult in analyzing and difficult to understand. Data needs to be transformed in order to obtain more correctness, increasing its use in data mining. In survey, we see many different approaches to solve these problems. In the data gathering phase, labeling is one of thecritical tasks to perform. Video cameras are being used to analyze road situations and a subtitle editor to find irregularities. Speed is another factor to be studied, as GPS signals can be non regular. The routes also need to be well defined as to avoid more traffic and not only due to safety norms, but also to prevent data quality from being compromised by these disturbances. The diversity of vehicles, drivers and recording devices are important that need to be taken for preventing over fitting.

3. Proposed Approach for Detecting Road Anomalies

For dealing with the difficulties discussed earlier, we propose a data mining approach. As data mining algorithms are able to find patterns in data, this can handle problems of hardware and speed impact, for example. At last it is expected to have better results than threshold based approaches, where these problems have a impact. For implementing the data mining approach, we will use CRISP-DM20 framework to guide us in all process. CRISP-DM is dividing in six stages. Each stage is divided in small operations in order to deal with a small set of problems. At the end of each stage, depending of the generated outcome, we execute another stage.

While studying the problems,it is observed that the most common irregularities are potholes, manholes and bumps. These three stages can categorize most of the irregularities found in our daily routine. Along with that we add one more stage i.e.; other: where we combine all irregularities that do not fit in the previous three stages. Thus, we have: Potholes – are mainly caused by bad quality or below surface issues, Manholes – are holes that allow workers access to some underground points. Generally allow access to infrastructures; Bumps normally man made, and usually used to slow down vehicles in crosswalk proximity; others irregularities that do not fit in any of the above mentioned stages. These are irregularities without any specific format. These canbe depressions or asphalt elevations. For this approach and as a output of road conditions in our area, we decided to focus only in four irregularities:manholes, long and short bumps, and others, leaving behind potholes. The reason for this is that there are few potholes in our area, so besides being difficult to acquire representative data from them, they do not represent typical irregularities. Also, we decided to divide bump types in two different categories aslong and short bumps. This division results from the possible danger level of each type. Long bumps are more harmful then short bumps; also, there are amore number of long and short bumps in our area.

As mentioned above, we will try to detect irregularities using Smartphone as data collecting device. We could use only 3-axis accelerometer to detect irregularities, but we need GPS coordinates to find out irregularitieslocation. Speed is relevant due to its influence in accelerometer data. For example, more speed causes high values in accelerometer data, which in turn can have a relation with the calculation of irregularities. We also collect timestamps to maintain irregularities history. This information could be used, for detecting irregularities, showing the evolution of the road situations. In this way, it creates a service to be used for drivers and government bodies. Here, all collected data are collected by Smartphone.Irregularities: string with the irregularitiestag or blank attribute. All attributes may have null data due to sensor errors; for example, data from GPS can be null due to improper GPS connection.

For collecting data, we developed an Android application that allows us to collect data from the sensors above described. Besides that, the system takes 50Hz sample rate, because, assuming 30cm as irregularities size, for a 50Km/h velocity (maximum speed allowed by law in city) we need a little less than 50Hz. In this application, we can label irregularitiessimultaneously with data collection from Smartphone sensors. As any touch in a Smartphone could change acceleration values, we used headphones with buttons, where each button corresponds to one irregularities type. In this way, we avoid touching the Smartphone. Labelingirregularities. All data is stored in *.csv* files. These *.csv* files are uploaded when the Smartphone has Internet connection. The data acquisition process was made with 3 different cars and 2 different Smartphone, in various days. To avoid external disturbances, most data collecting occurred during the night time and in low traffic zones. Thus, the process is safer and we avoid other problems like traffic jams and crossing roads, which would reduce the data quality.

Collected Irregularities as per survey:

Irregularities	Number
Manhole	36
Short Bump	53
Long Bump	51
Other	17
Total	157

Data Preparationafter data collection, we started the data preparation stage. In this step, we need to clean data, i.e. nullvalues, deal with possible inconsistent data, extract features, and transform data as represented in following steps.

Step 1: CSV Files
Step 2: Collecting Lables
Step 3: Null Elimination
Step 4: Delete inconsistent data
Step 5: Features Extraction
Step 6: CSV Files

The labeling of data was carried out simultaneously with the driving part. In this specific case, as we use a 50 Hz sample rate in the sensors, it was impossible to label the exact data. So, it was required to review the complete data and correct the labeling. This was a manual process to see where the irregularity started and ended. With this data, it was possible to fill the irregularitylabels in all data.

ISSN: 2454-4248
719 – 721

Analyzing the difficulties problem and ours goals, considering that we have a solid ground truth with irregularity correctly labeled, we face a difficulty of classification. For this we can use supervised learning algorithms, such as decisions trees, for example. This way, in order to choose the best algorithm for this problem and study possible alternatives, we perform the certain tasks. First, we use the algorithms with the standard pre-defined values. After analyzing all the results, some changes can be made in each algorithm parameters looking for better outputs, bearing in mind that tweaking the algorithm can cause model overfitting. At this moment, we anticipate that tree-based models (like Random Forest (RF), Gradient Boosting (GB) or Decisions Trees (DT)) can accomplish good results for this problem. Neural Networks (NN) can also accomplish good outputs, but can have performance problems in training, depending of data size. Support Vector Machines (SVM) have the same training performance problems as NN.

4. Conclusions and Future Work

The target of this system is detecting road irregularity using Smartphone sensors and data mining. Data was labeled and categorizedinto five classes. We obtained features, creating new attributes and using information already present. Due to classification problems, we tested different algorithms. As we have a dataset with several attributes that can be performed over fitting and noise, decreasing algorithms performance, Weka tool was used for obtaining the best attributes. We also run the same algorithms on new dataset, and observed an improvement. In order to obtain better outputs, we can use other strategies.

References

[1] Vittorio A, Rosolino V, Teresa I, Vittoria CM, Vincenzo PG, Francesco DM. Automated Sensing System for Monitoring of Road Surface Quality by Mobile Devices. Procedia - Soc Behav Sci. 2014;111:242-251. doi:10.1016/j.sbspro.2014.01.057.

[2] Douangphachanh V, Oneyama H. Exploring the Use of Smartphone Accelerometer and Gyroscope to Study on the Estimation of Road Surface Roughness Condition. In: Proceedings of the 11th International Conference on Informatics in Control, Automation and Robotics. SCITEPRESS - Science and and Technology Publications; 2014:783-787. doi:10.5220/0005117407830787.

[3] Douangphachanh V, Oneyama H, Engineering E. A Study on the Use of Smartphones for Road Roughness Condition Estimation. Proc East Asia Soc Transp Stud. 2013;9(2007):14. doi:10.1186/1687-1499-2014-114.

[4] Nomura T, Shiraishi Y. A method for estimating road surface conditions with a smartphone. Int J Informatics Soc. 2015;7(1):29-36.

[5] Chen K, Tan G, Lu M, Wu J. CRSM: a practical crowdsourcing-based road surface monitoring system. Wirel Networks. 2016;22(3):765-779. doi:10.1007/s11276-015-0996-y.

[6] Kalra N, Chugh G, Bansal D. Analyzing Driving and Road Events via Smartphone. Int J Comput Appl. 2014;98(12):5-9. doi:10.5120/17233-7561.

[7] Mohan P, Padmanabhan VN, Ramjee R. Nericell: Rich Monitoring of Road and Traffic Conditions using Mobile Smartphones. Proc 6th ACM Conf Embed Netw Sens Syst - SenSys ’08. 2008:323. doi:10.1145/1460412.1460444.

[8] Bo C, Jung T, Mao X, Li X, Wang Y. SmartLoc: sensing landmarks silently for smartphone-based metropolitan localization. EURASIP J Wirel Commun Netw. 2016;2016(1):111. doi:10.1186/s13638-016-0603-7.

[9] Gps Error Analysis. In: Global Positioning System: Theory and Applications, Volume I. Washington DC: American Institute of Aeronautics and Astronautics; 1996:469-483. doi:10.2514/5.9781600866388.0469.0483.

[10] Zeng H, Park H, Fontaine MD, Smith BL, McGhee KK. Identifying Deficient Pavement Sections by Means of an Improved Acceleration-Based Metric. Transp Res Rec J Transp Res Board. 2015;2523:133-142. doi:10.3141/2523-15.

[11] Seraj F, Zhang K, Turkes O, Meratnia N, Havinga PJM. A smartphone based method to enhance road pavement anomaly detection by analyzing the driver behavior. In: Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2015 ACM International Symposium on Wearable Computers - UbiComp ’15. New York, New York, USA: ACM Press; 2015:1169-1177. doi:10.1145/2800835.2800981. 12. Mukherjee A, Majhi S. Characterisation of road bumps using smartphones. Eur Transp Res Rev. 2016;8(2):13. doi:10.1007/s12544-016-0200-1.

[12] Douangphachanh V, Oneyama H. Formulation of a simple model to estimate road surface roughness condition from Android smartphone sensors. IEEE ISSNIP 2014 - 2014 IEEE 9th Int Conf Intell Sensors, Sens Networks Inf Process Conf Proc. 2014;(April):21-24. doi:10.1109/ISSNIP.2014.6827694.

[13] Chen S-Y, Shih A, Hsiao C-Y. Road conditions detection using Arduino based sensing module and smartphone. In: 2015 IEEE International Conference on Consumer Electronics - Taiwan. IEEE; 2015:254-255. doi:10.1109/ICCE-TW.2015.7216884.

[14] Bello-Salau H, Aibinu A, Onwuka E, et al. A New Measure for Analysing Accelerometer Data towards Developing Efficient Road Defect Profiling Systems. J Sci Res Reports. 2015;7(2):108-116. doi:10.9734/JSRR/2015/16840.