A Crowdsourcing-based Road Anomaly Classification System

Ru-Yu Wang Yi-Ta Chuang Chih-Wei Yi
Department of Computer Science, National Chiao Tung University
1001 University Road, Hsinchu City 30010, Taiwan
izeki.cs96g@nctu.edu.tw

***Abstract*—Road networks are the most important facility to the public transportation in modern cities. Governments around the world allocate large amounts of budgets for the pavement maintenance every year. In this paper, we proposed a crowdsourcing solution to categorize road anomalies into safety related anomalies such as speed bumps and rumble strips, and dangerous anomalies such as bumps and potholes. The proposed system is composed of three parts: a smart probe car crowds (SPC-crowd) that serve as the anomaly data source; cloud servers that are the core for the anomaly classification; and application services that provide various innovative applications to facilitate the pavement maintenance. To support the crowdsourcing procedure, in the SPC-crowd side, we proposed cross-SPC techniques by adopting the underdamped oscillation model (UOM). In the cloud side, a supervised learning classification model was adopted on the anomaly data generated from the SPC-crowd. To validate the proposed system, extensive field trial was performed. The experimental results shown that our system can facilitate the pavement maintenance through the crowdsourcing solution.**

***Index Terms*—Smart probe car, crowdsourcing, data mining, support vector machine**

I. Introduction

Road surface conditions have been a public issue in modern society. It affects not only the travel experience but also the driving safety. According to the report from the Ministry of Transportation and Communications, Taiwan, the governments have paid several hundred million NT dollars on national compensation for the accidents caused by poor road surfaces during last ten years. However, the performance of the project has always been criticized by the citizens. It is not only because of the expensive manpower, but also for the inefficient monitoring by humans. How to maintain road surface in an efficient and intelligent way is an important issue.

Our proposed system consists of three subsystems: the smart probe car (SPC)-crowd subsystem, the server subsystem and the application subsystem. The SPC-crowd subsystem comprises a crowd of SPCs. An SPC is the dynamic couple of an ordinary vehicle and a smartphone. The smartphone mounted on the vehicle executes sensing program to detect bumping events caused by the vehicle running over road anomalies. SPCs report detected abnormal events along with feature data to the server subsystem via wireless Internet connections for further process. The server subsystem collects data from the SPC crowd, and then classifies detected events into two categories: dangerous anomalies and safety related anomalies. The former includes potholes and bumps which are caused by the dynamics of weather and constant stresses of traffic and are needed to be repaired for public safety concerns; the latter are speed bumps and rumble strips which are used to alert motorists to slow down. By utilizing the road anomaly information, the application subsystem can provide various advanced services. For example, the dangerous anomaly information is helpful for motorists to pay attention on dangerous road condition ahead. The dangerous anomaly information is also useful for the government to arrange long-term maintenances.

Many solutions based on mobile sensing technologies have been proposed to detect road anomalies in the past few years. In the perspective of detection methods, it can be categorized to detection by vehicle vibrations [1], [2], [3], [4], [5], [6], [7], image processing [8], [9] and ultrasonic [10]. However, to obtain high accurate results, an expensive measurement device is required. Recently, the popularity of smartphones make a perfect candidate on the application. Nowadays, accelerometer and GPS are the basic equipment for the smartphone. In addition, the increasing computing power, storage and rich wireless connectivity make smartphones possible to cope with the road anomaly detection problem both in low cost and crowdsourcing manner.

Note that most of literature works only focus on the detection of road anomalies and seldom talk about anomaly classification. The safety related anomalies such as speed bumps and ramble strips are used to alert dangerous road conditions ahead. In the point of view from road pavement managers, the safety related anomalies should not be treated as dangerous anomalies.

In this paper, our proposed system has several features. First, the crowdsourcing concept is achieved by the SPC crowd. The SPC-crowd as the basic entity of the system provides a low cost and large scale solution to support the data source of the system. Second, to support cross-SPC road anomaly detection, the anomaly ranking mechanism based on the underdamped oscillation model (UOM) is adopted [5]. Finally, by the crowdsourced data from the SPC-crowd, the server subsystem is able to classify the road anomalies into dangerous anomalies and safety related anomalies by adopting the machine learning tool. In summary, our system improve the efficiency of automatic road quality monitoring.

II. Related Works

For methods by vehicle vibrations, [1] proposed a public transport network which makes use of buses as the carrier of sensors to sensing road anomalies on the road surface. In [2], the pothole patrol system was proposed with five filters to cope with the speed, sudden acceleration, veering, braking, subtle changes of sensor orientation and horizontal movements. However, the system only considers a fixed sensor orientation scenario. In [5], [6], [7] combine the smartphone and the vehicle into a smart probe car. It's the most popular method adopted in the literature. And in our paper, the detection method is based on [7].

For methods by image processing, it requires a camera installed on vehicles and capture images of road surfaces [8], [9]. The limitation of this kind of methods is that the quality of detection result affected significantly by the angle of the camera and the weather.

For methods by ultrasonic, an ultrasonic transducer is equipped on vehicles [10], the ultrasonic waves are continuously emitted to the road surfaces and the anomalies are detected by the reflection time. However, to obtain high accurate results, an expensive measurement device is required.

III. Crowdsourcing-based Road Anoamly Classification System

A. System Architecture

As illustrated in Fig. 1, the proposed system comprise three subsystems.

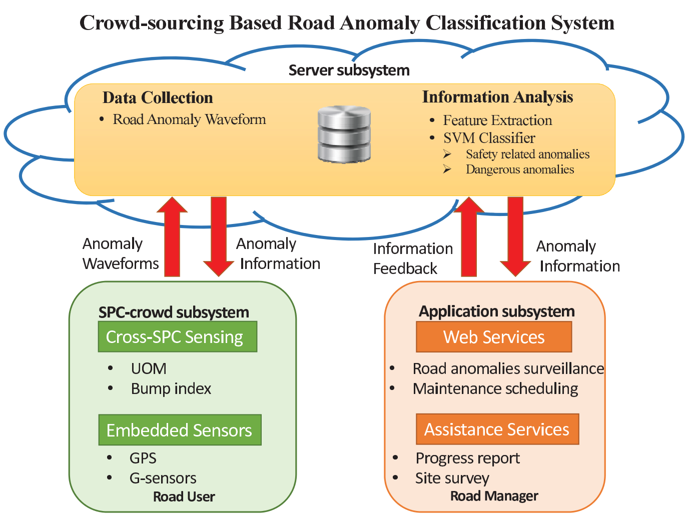

Fig. 1. The system architecture.

For the SPC-crowd subsystem, the vehicle embedded with GPS and G-sensors keeps sensing the vibrations when the vehicle is moving. And the road anomalies detection algorithm is based on the Underdamped Oscillation Systems (UOM), which is used to model the waveform during the duration of the vibrations. Then the bump index (BI) can be calculated from the ratio of the standard deviation of the anomalistic road and the normal road which can be deduced from the UOM. For the sever subsystem, the features will first be extract form the raw data, then the support vector machines (SVM) is used to classify the detected road anomalies into safety related anomalies and dangerous anomalies. At last, for the application, web services and assistance services will be provided to the users.

B. Processing flow

Figure 2 illustrates the processing flow of the system.

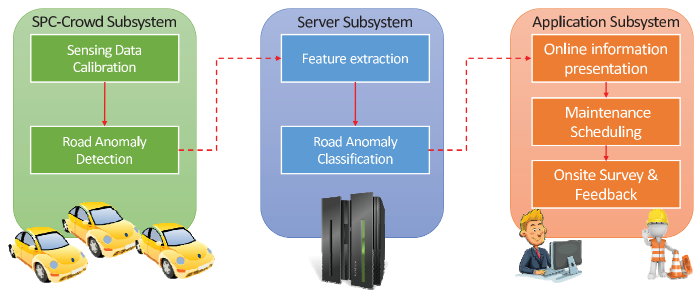

Fig. 2. The processing flow of the system.

The workflow of our proposed system is showing as follow: first, in the SPC-crowd subsystem, we attach the smartphone in the front of the vehicle, prevent it from the unnecessary vibration. And we run the application which will first calibrate the three-axis of the smartphone, make sure that x-axis points to the front of the vehicle, y-axis points to the right of the vehicle and z-axis points to the gravity. Then it starts to collect vibration from the vehicle, and detects road anomalies. Once a anomaly is detected, the features of the anomaly is recorded. If the wireless connection is available, the recorded data will be sent back to server subsystem for further analysis. In the server subsystem, the server collect data from all the SPCs and use Support vector machine (SVM) as classifier, classify the detected anomalies into dangerous anomalies and safety related anomalies. The server subsystem will store the result of the classification in the database for further usage. Finally, in the application subsystem, we provide a website which plot all the road anomalies on the Google Map, and user can easily search the road and see how many pothole on the road surface.

IV. SPC-Level Anomaly Detection

A. Vertical Component Extraction

We observe that the gravity is the major component of the g-vectors at most time. Therefore, in case the smartphone is mounted on the rack, by applying principle component analysis on a collection of g-vectors, we can find the direction of the gravity.

Let $\overline{\mathbf{g}}_t$ denote the running average of the g-vector by time t over a time period T_1. The calculation of $\overline{\mathbf{g}}_t$ can be expressed as

$$\overline{\mathbf{g}}_t = - \underset{(t-T_1 \le s < t) \wedge (\left\| \mathbf{g}_s - \overline{\mathbf{g}}_{t-} \right\| \le \delta)}{\operatorname{avg}} \mathbf{g}_s \tag{1}$$

where δ is a stable threshold to filter out contaminated g-vectors that are significantly different from the previous running average. In our implementation, δ is set to $2m/s^2$. Let $g_t^{\perp}$ denote the VC of the g-vector at time t. Then, we have

$$g_t^{\perp} = \langle \mathbf{g}_t, \overline{\mathbf{g}}_t \rangle / \|\overline{\mathbf{g}}_t\| . \tag{2}$$

B. Bumping Index

1) Underdamped Oscillation System: An underdamped oscillation system, called the Kelvin model, where m, x, v and a respectively denotes the mass, displacement, velocity, and acceleration of the object, k is the spring constant of the string, b is the damping coefficient of the damper, and $f = -kx - bv$ is the force applied to the object. The deduction details is showed in [7].

Consider a vehicle running on a road. The springs connecting to tires are compressed as the vehicle hits the anomaly; the tires then oscillate like a damped system. In this work, we would like to model the waveform during the duration of the vibrations by the UOM.

2) Anomaly Event Detection: According to the model $a(t) = A\omega^2 \cos(\omega t) e^{-\lambda t}$, the standard deviation of the acceleration from $t = t_0$ to $t = t_1$ can be calculated by

$$\sigma = A\sqrt{\frac{1}{t_1 - t_0}\left(\begin{array}{c} \int_{t_1}^{t_0} \left(e^{-\lambda t}\cos\left(\omega t - \theta_0\right)\right)^2 dt \\ -\left(\int_{t_1}^{t_0} e^{-\lambda t}\cos\left(\omega t - \theta_0\right) dt\right)^2 \end{array}\right)}. \tag{3}$$

From Eq. 3, we observe that the standard deviation is proportional to the initial amplitude parameter A and an integration of a function of cosine over the duration. The square root part can be viewed as a constant if we fixed the duration and selected the same θ_0. Then the standard deviation is only related to the amplitude. To detect the bumping event, we divide the standard deviation during the bumping periods by the standard deviation during the normal period.

The ratio can be a cross-SPC index on the road anomalies because the only factor is the amplitude of the vibration. There are three steps in our detecting algorithm:

1) Divide the data into different speed sets. Each set has different threshold for detecting. We first identify if the incoming data may be a suspect of anomaly.

$$VI = \frac{\Delta g_i^{\perp}}{\sigma_{normal}} > T_1 \tag{4}$$

Where VI is the first index to decide if there is a road anomaly, $\Delta g_i^{\perp}$ is the maximum change of z-axis value in one second, σ_{normal} is the standard deviation of normal road, and T_1 is the threshold in this step.

2) Then we will cut a 0.5 second vibration wave start from the point which is over the threshold, and calculate the standard deviation of this interval.

$$BI = \frac{\sigma_{event}}{\sigma_{normal}} > T_2 \tag{5}$$

Where BI is the index which indicate a road anomaly, and T_2 is the threshold in this step.

3) If BI exceed the threshold T_2, then we mark it as a road anomaly and record the location and the timestamp.

V. System-Level Anomaly Classification

A. Feature Extraction

After detection, we need to extract features form the waveform for further classification. According to the observation, the waveform of anomaly events have similar pattern: each waveform consist of a first shock and an aftershock. In Fig. 3 , the first shock is caused by the front wheel running over the road anomaly, while the aftershock is caused by the rear wheel. Based on the observation, we pick up the following feature to train our SVM.

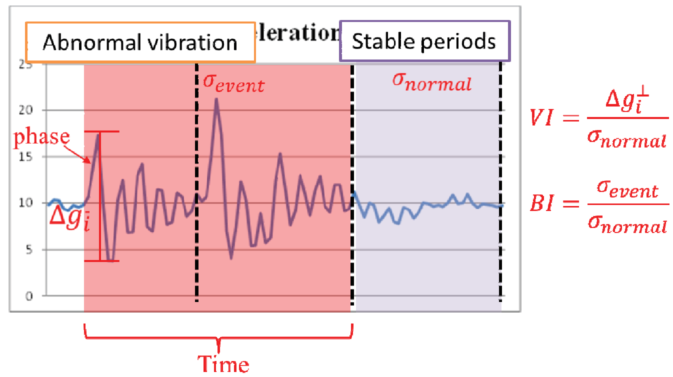

Fig. 3. The waveform of an road anomaly.

1) Phase: if the phase at the start point rise up, the road anomaly comes out to be a bump. Otherwise, we consider it as a pothole.
2) Time: The duration of the road anomaly, which can be the index of the severity.
3) BI: the ratio of the standard deviation of anomaly event and the standard deviation of normal road.
4) VI: the ratio of the peak of the anomaly and the standard deviation of normal road.
5) Amplitude: the peak value of the waveform of the road anomaly.

B. Anomaly Classification

Support vector machines are supervised learning models with associated learning algorithms that analyze data used for classification and regression analysis. The principle of SVM is finding hyper-planes in R^d space which could separate data into different groups. The formulas which need to be solved are showing as following:

$$\min_{w,b} \frac{1}{2} w^T w \tag{6}$$

$$y_i \left(\left(w^T x_i \right) + b \right) \geq 1 \tag{7}$$

Where x_i , y_i, $i = \{1, ..., n\}$ is the collection of data set, $x_i \in R^d$, $y_i \in \{1, -1\}$ and w is normal vector of the plane.

In our case, SVM is a suitable tool to classification road anomalies into dangerous anomalies and safety related anomalies.

VI. Experimental Results

A. Experimental Setup

To validate our proposed system, we conducted a field experiment in real-world driving environments. We took NCTU campus as our testing environment, the route is showing in fig.. We ran five times on this route, and each route is 3.5km long with about 25 road anomalies. Finally, we have data of 17.5km long route and 119 detected road anomalies. The experiment equipment sets up is shown below: we use two smart phones to record data. One is Sony Xperia Z3, which is attached on the front board of the vehicle and collecting the vibration data while the vehicle is moving. The sampling rate of the Sony Xperia Z3 is about 198Hz. The other one is ASUS Padfone, we attached it on the license plate and used it to film the ground in front of the vehicle, so we can know what kinds of road anomalies the vehicle run through during the experiment. The vehicle we rent is Nissan Livida.

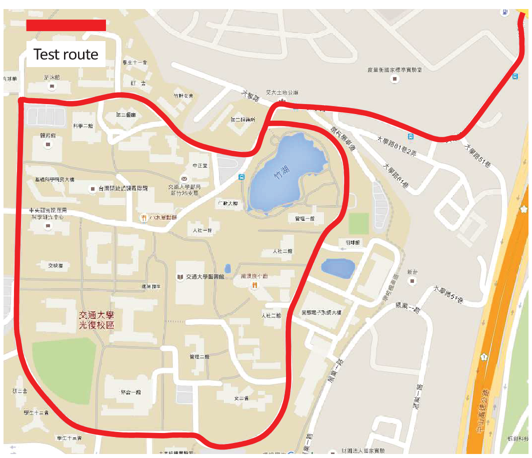

Fig. 4. The test field in NCTU.

B. Experimental Results

After the experiment, we process the collecting data as we described in sec.III. We use ten-folds cross validation to test and verify our data, the result of classification is shown in Fig.. The precision of safety related anomalies is about 88% and the precision of dangerous anomalies is about 82%. The error rate of each class is 12% and 18%, it is because of the pattern of safety related anomalies and dangerous anomalies may be similar at times, the way to deal with this problem is collect much more data. In the future work, we are going to collecting more data for the experiment by crowd-sourcing and improve the performance of our proposed system.

VII. Conclusions

In this paper, we presented a crowdsourcing-based road anoamly classification system. The system is composed of SPC-crowd subsystem, server subsystem and application subsystem. We collect crowd-sourcing data from the SPC-crowd

		Actual type	
		Safety related	Dangerous
Predict type	Safety related	60	9
	Dangerous	8	42
Total number of anomalies		68	51
Precision		0.88	0.82

Fig. 5. The result of the experiment.

subsystem, process and classify the data on the server subsystem and show the road information to users at the end. The experimental results show that the precision of the classification on safety related anomalies is 88% and on dangerous is 82%. In the future, we will improve the performance of our proposed system and apply real-world crowd-sourcing data into the experiment.

Acknowledgement

This work of C.-W. Yi was supported in part by the National Science Council under Grant MOST 104-2221-E-009-026-MY2, MOST 104-3115-E-009-002, MOST 105-2218-E-009-003, and MOST 105-2221-E-009-102-MY3; by Sinica under Grant AS-105-TP-A07; and by ITRI under the IoT/M2M Wearable-based Healthcare for Active Aging - Activity/Calorie Monitoring and Analysis at Home (2) Project.

References

[1] K. D. Zoysa, C. Keppitiyagama, G. P. Seneviratne, and W. W. A. T. Shihan, "A public transport system based sensor network for road surface condition monitoring," in *ACM SIGCOMM workshop on Networked Systems for Developing Regions (NSDR '07)*, 27 August 2007, pp. 91–96.

[2] J. Eriksson, L. Girod, B. Hull, R. Newron, S. Madden, and H. Balakrishnan, "The pothole patrol: Using a mobile sensor network for road surface monitoring," in *the 6th International Conference on Mobile Systems, Applications and Services (MobiSys 08)*, 17-20 June 2008, pp. 29–39.

[3] P. Mohan, V. N. Padmanabhan, and R. Ramjee, "Nericell: Rich monitoring of road and traffic conditions using mobile smartphones," in *the 6th International Conference on Embedded Network Sensor Systems (Sensys 08)*, 5-7 November 2008, pp. 357–358.

[4] A. Mednis, G. Strazdins, R. Zviedris, G. Kanonirs, and L. Selavo, "Real time pothole detection using android smartphones with accelerometers," in *2011 International Conference on Distributed Computing in Sensor Systems and Workshops (DCOSS)*, 27-29 June 2011, pp. 1–6.

[5] Y.-T. Chuang, C.-W. Yi, T.-H. Wu, and C.-Y. Chan, "Model the vibrations of rack-mounted on-vehicle devices by underdamped oscillations," in *the 17th International Conference on Intelligent Transportation Systems (ITSC)*, 8-11 October 2014, pp. 982–987.

[6] K. Chen, G. Tan, M. Lu, and J. Wu, "Crsm: a practical crowdsourcing-based road surface monitoring system," *Wireless Networks*, pp. 1–15, June 2015.

[7] C.-W. Yi, Y.-T. Chuang, and C.-S. Nian, "Toward crowdsourcing-based road pavement monitoring by mobile sensing technologies," *Intelligent Transportation Systems*, vol. 16, no. 4, pp. 1905–1917, April 2015.

[8] J. Lin and Y. Liu, "Potholes detection based on svm in the pavement distress image," in *2010 Ninth International Symposium on Distributed Computing and Applications to Business Engineering and Science (DCABES)*, 10-12 August 2010, pp. 544–547.

[9] X. Yu and X. Yu, "Pavement pothole detection and severity measurement using laser imaging," in *2011 IEEE International Conference on Electro/Information Technology (EIT)*, 15-17 May 2011.

[10] R. Madli, S. Hebbar, . Pattar, and V. Golla, "Automatic detection and notification of potholes and humps on roads to aid drivers," *IEEE Sensors Journal*, vol. 5, no. 8, pp. 4313–4318, March 2015.