MagTrack: Detecting Road Surface Condition using Smartphone Sensors and Machine Learning

Meenu Rani Dey
Information Security Lab
IIIT Bhubaneswar
Odisha, India
a117004@iiit-bh.ac.in

Utkalika Satapathy
Information Security Lab
IIIT Bhubaneswar
Odisha, India
a117010@iiit-bh.ac.in

Pranali Bhanse
Information Security Lab
IIIT Bhubaneswar
Odisha, India
a117002@iiit-bh.ac.in

Bhabendu Kr. Mohanta
Information Security Lab
IIIT Bhubaneswar
Odisha, India
c116004@iiit-bh.ac.in

Debasish Jena
Information Security Lab
IIIT Bhubaneswar
Odisha, India
debasish@iiit-bh.ac.in

***Abstract*—Low maintenance of the roads is one of the extensive cause of increasing road accidents and vehicle breakage. Mostly roads contain potholes, wreckage or detritus which is best to be bypassed however if not sometimes lead to severe road causalities. Many car or bus accident happens over a wrecked bridge, as it slips and overturns. To avoid this kind of mishap and improve the safety of the road we have proposed a machine learning based model, MagTrack, to detect these road conditions and inform apriori to the responsible authority for fast reparation as well as other passers-by to drive carefully or take an alternative route. We have collected various road surface conditions such as smooth roads, uneven roads, potholes, speed breakers, and rumble strips data using magnetometer and accelerometer sensor embedded in our Smartphone and analyzed using various classification algorithms like Random Forest (RF), Random Tree (RT) and Support Vector Machine (SVM). The classification has done after performing the feature selection using GreedyStepwise, Ranker, and BestFirst Algorithms considering minimum, maximum, median and standard deviation as statistical features. 92% of accuracy to detect the road surface condition has been achieved by MagTrack.**

***Index Terms*—Smartphone, Magnetometer & Accelerometer, Machine Learning, Road Surface Condition, Vehicles.**

I. Introduction

In today's world, Roads and vehicles traffic is kept an important part of day-to-day lives thereby, road surface condition is the main concern nowadays. Therefore, monitoring the road surface condition is a significantly most important aspect and takes a lot of attention, ensuring the comfort and safety of the user's. Also, roads play a very necessary contribution to economic development, growth and important for social benefits. As a solution, many methods/applications are proposed. Nowadays, smartphones gaining a lot of attention because of various sensors equipped in them like microphones, camera, GPS, accelerometer, magnetometer, etc. So, built-in sensors can be used to monitor the road condition, traffic analysis, route information and speed, etc. The prior work has been done in this field to develop road, detect traffic monitoring [1], analyzing driving behavior [2] and deployed devoted sensors on a roadside and/or on vehicles [3] that are most expensive and required lot of human efforts. Generally, road surface condition is defined by the irregularity affected by the bad quality ride of the vehicles. Bad road condition increases the number of accidents, damages vehicles also consume more fuel, thereby it increases the maintenance cost of the user's vehicles.

El-Wakeel et al. [4] demonstrates and analyze different types of road surface condition by the linear acceleration of vehicles using an inertial sensor called MEMS mounted inside two different types of vehicles and land. Sharma et al. [5] designed an "S-Road Assist" application for smartphones that tracks the driving behavior of drivers. Mohan et al. [6] presents a system called Nericell for sensing road and traffic condition using piggybacking on smartphones, that focuses some sensing components that use the GSM radio, microphone, accelerometer and GPS sensors which are located inside these smartphones to detect bumps, braking, potholes, and honking. White et al. [7] described how smartphones can automatically identify the traffic accident through accelerometers and acoustic data, after the accident it also notifies to the emergency server about the situation using the photograph. Collotta et al. [8] has given a road traffic management approach using a wireless sensor network and uses a magnetic sensor, Further, they use a fuzzy logic controller that manages the workload of the network. Liu et al. [9] describes leaser scanning for muddy and water surface detection using leaser range for various conditions, also he described weak and strong mirror deflection phenomena. Zhang et al [10] proposed a traffic monitoring scheme that preserve privacy using speed of the vehicles and information of different routes. To ensure the privacy of vehicle's speeds they use homomorphic Paillier cryptosystem . Mohamed et al. [11] design a framework for detecting road speed bumps using smartphone he also identifies to detect speed bumps they use gyroscope around gravity rotation, in

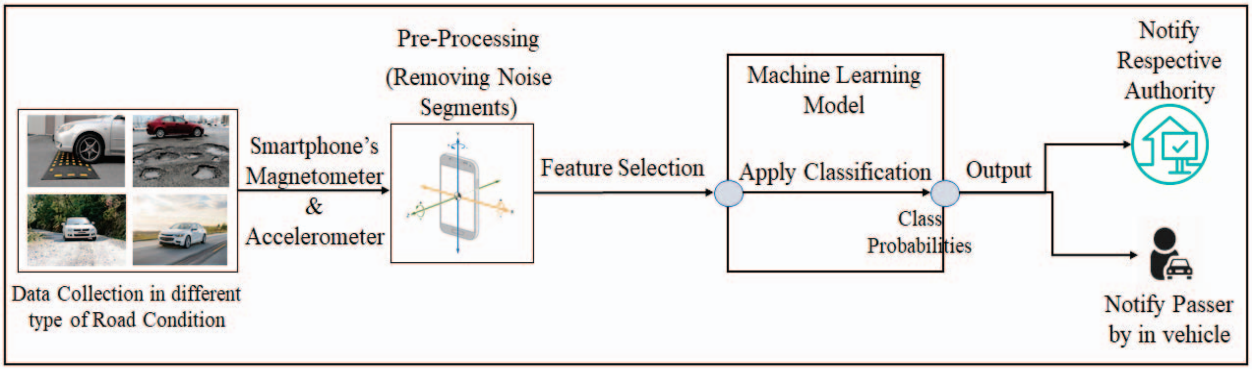


Fig. 1: Magtrack's System Architecture

addition to the accelerometer as cross-validation.

We have proposed a new model "MagTrack" that monitors road surface condition through Machine Learning using a smartphone.

The paper is organized as follows: Section 2 Discuss the background needed for the proposed model. Section 3 describes the system architecture. The experimental setup is described in section 4. Design and implementation are done in section 5. Section 6 describes the result of the proposed method. Finally, section 7 summarized the conclusion and future work.

II. Background

A. Feature Extraction Algorithm:

1) Ranker Algorithm: Ranker, evaluate the individual feature and according to that, it ranks the features. Feature which has higher rank will be considered for classification and which have lower rank will be discarded.

2) GreedyStepwise Algorithm: The GreedyStepwise algorithm works on space and features according to this, it will search forward and backward and gives those features which have more impact on classification.

3) BestFirst Algorthim: In BestFirst Search start with an empty set of attribute and traverse forward or starts with full set of attribute and traverse backward or starts with mid of the set and traverse both the directions.

B. Classification Algorithm:

1) SVM Algorithm:: SVM handles a large amount of data using less mathematical computation. The main idea behind SVM is: The algorithm which separates the classes through a line named as Decision Boundary The goal of Decision Boundary is maximizing the margin between the points which are placed on either side of Decision Boundary [12].

2) RandomTree Algorithm: RandomTree works exactly like a decision tree for tree-based classification.

3) RandomForest Algorithm: RandomForest is easy and flexible ML algorithm which can be used for both Regression and classification, and can get the most efficient result without hyper-parameter tuning. For better accuracy, Random Forest constructs more then one decision tree and combines together and then try to classify the input features. [13].

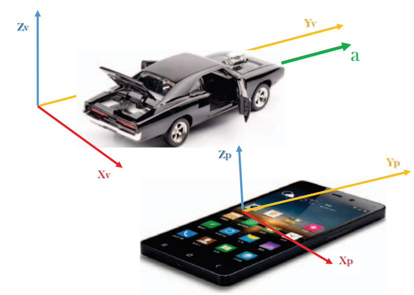


Fig. 2: Properly aligned coordinate system of Phone and Vehicle

III. System Architecture

The main components of our proposed model, MagTrack are smartphone to collect the data and weka tool for running ML algorithms as shown in Fig. 1. Basically MagTrack includes 7 stages as shown in Fig. 3. In Phase 1, the magnetometer and accelerometer of the smartphone are initialized followed by data collection in various road conditions (smooth, uneven, potholes and rumble strips etc) in phase 2. Noises are removed from the collected data in phase 3. From the noise free data, features are extracted using BestFit, Ranker and GreedyStepp-wise algorithm in phase 4. Phase 5 takes the feature extracted data as input and using Classification Algorithm classifies the data as various classes. The output class is analysed in phase 6 and finally in the last phase action is taken based on the output.

IV. Experimental Setup

For MagTrack, we have used hardware and software tools for our work.

A. Hardware Setup (Smartphone)

For data collection we have used smartphone. A smartphone have so many capabilities related to motion, acceleration etc.

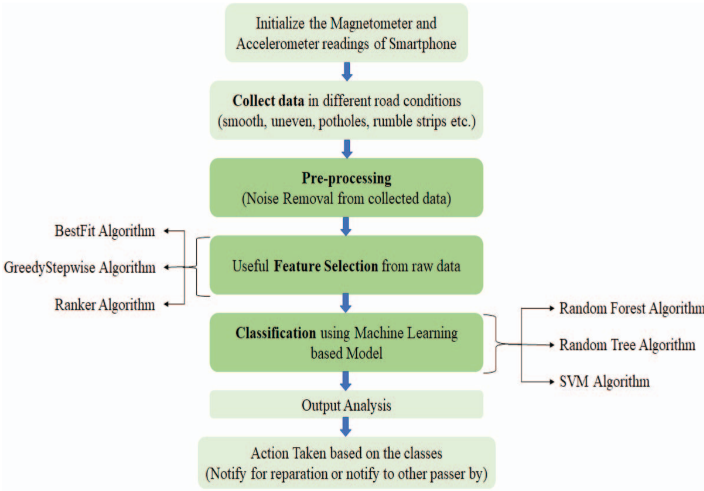


Fig. 3: Working flowchat of MagTrack

1) Sensing: Motion (Magnetometer, Accelerometer), Localization (GPS), Visual (Camera)

2) Communication: Cellular data (UMTS, GPRS, EDGE)

3) Computation: OS, CPU, storage. This all capabilities are exist in most of the smartphone in a single package, which are available in market. In our work we have used Redmi note 5 with 4GB RAM and snapdragon 625 processor and also MotoG 5S with same specification.

B. Software Setup

Weka is a ML software tool used for knowledge analysis. It provides data visualization, data preparation, and other models and algorithms like data association, data clustering, data regression, data classification and rules of mining.

V. Design and Implementation

A. Data Collection

For MagTrack, we have developed one mobile app to collect data from different sensors present in smartphones like magnetometer and accelerometer. MagTrack application supports various modes like driving, cycling, walking and according to the mode, it automatically collects the data. Driving mode detection algorithm takes accelerometer values for 3-axes and then finds RMS (root mean square). For more accurate GPS data it gives information about the position of a vehicle. To achieve an accurate result and better conclusion, the quality and quantity of the data plays a very important role in our model. Therefore, for analysis of road surface conditions, we have tried

$$7R \times 5\,L \times 5\,RC = 245\,Sets \tag{1}$$

Where R, C, L, RC indicates Readings per Condition, Location, Road Condition respectively.
The 7 road surface conditions are as follows: 1) Idle 2) Uneven road 3) Pothole 4) Rumble strip 5) Speed hump and 6) Smooth road We have collected data using multiple smartphones for various road surface conditions at different locations. While collecting data, the vehicle maintained an average speed of 20 km/hr. In total we have covered approximately 100km. For providing veracity to MagTrack we placed smartphone inside the vehicle at different location like dashboard, cupboard, dock, and inside the pocket while driving. The main objective of MagTrack is to collect these information in every possible situation so that the simple algorithms with large amount of data can easily deal with the complicated data analytics. The results of accelerator's sensor data are shown in Fig 4.

TABLE I: List of Parametres used

Classifier	Optimal Parameters
RandomForest	S = 1, f = 3, i = 100
RandomTree	S = 1, f = 3, i = 100
SVM	r = 0, C = 16, d = 1, $\gamma = 0:5$

B. Re-orientation

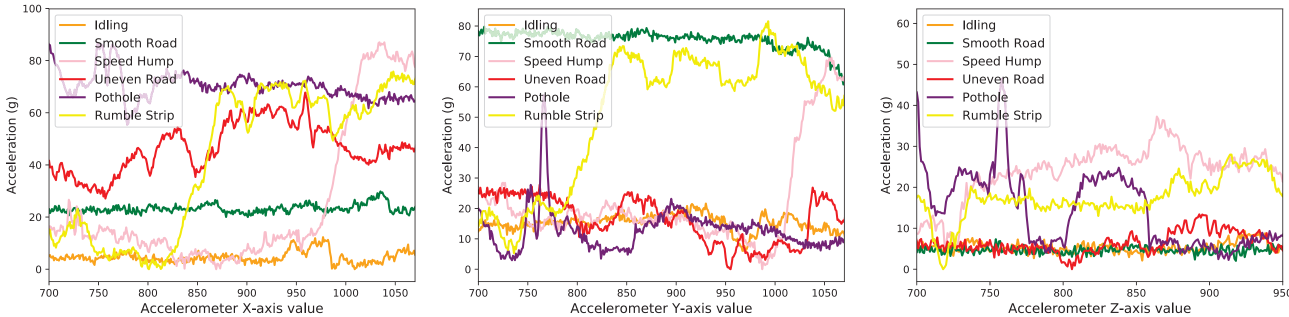
Though there is no physical connection between a smartphone sensor and vehicle, Magtrack detects the current road condition. The smartphone's accelerometer gives proper readings for all three directions when the orientation and alignment of smartphone and vehicle are the same. But, in real scenario the road condition may vary and also smartphone can be placed anywhere inside a vehicle for e.g. car seat, bag, clamps, dashboard, etc which changes it's orientation and alignment making non-trivial for accelerometer measurement. Hence, the axes of the phone should be reoriented with the axes of the vehicle as shown in Fig 2 [14]. After the reorientation of the axes, the acceleration reading predicts the road surface condition. The MagTrack Should have to perform continuous virtual reorientation because the acceleration and deceleration can be erroneous for different road condition like speed breakers, potholes, etc. For calculating the angle inclination we have followed one approach that in [6]. We have applied the reorientation algorithm in every 1 second. Reorienting sensor value also takes care of the impacts in sensor when an user is talking over phone for a long period of time.

C. Feature Extraction

The smartphone provides raw value of acceleration data in the unit of g-force and magnetic field data in the unit of microTesla. The raw accelerometer data further processed for extracting valuable features. Table II. summarizes the list of extracted feature. For optimizing the feature, we have applied feature selection algorithm named as BestFirst, GreedyStepwise, and Ranker. The optimized features further used for classification.

D. Classification

For classification, we have considered three ML algorithm named as SVM, RandomForest, RandomTree in MagTrack. For SVM we used radial basis function kernel with four parameters, degree (d), cost (C), coefficient (r), and gamma (Γ). On the other hand for RandomForest and RandomTree,



(a) (b) (c)

Fig. 4: Accelerometer Readings for Various Road Condition

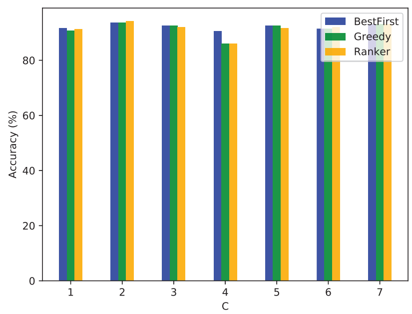


Fig. 5: SVM

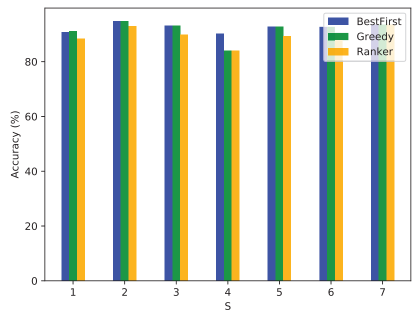


Fig. 6: RandomTree

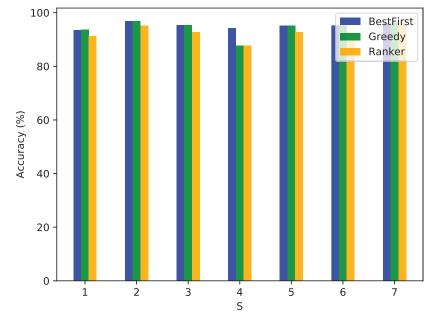


Fig. 7: RandomForest

TABLE II: Features used in classification

Domain	Parameters	Domain	Parameters
Time	Min Max Mean Median Standard Deviation Skewness Energy	Frequency	Min Max Mean Median Standard Deviation Skewness Energy

TABLE III: Comparison between different classifier and feature selection techniques Based on Accuracy

Feature Selection Techniques	Classifier		
	SVM	RandomForest	RandomTree
BestFirst	92.23	95.23	92.60
GreedyStepwise	91.45	94.37	91.76
Ranker	91.52	92.70	90.89
Average	91.73	94.10	91.75

we used three parameters, number of Iteration (i), number of Features (f), and seeds (S). All features values are listed in Table I. We have used Weka tool for evaluating the effectiveness of various features for categorizing the different road surface conditions. The data is normalized in the range [0,1] before applying the classification algorithm.

VI. Result and Discussion

A. Magnetometer and Accelerometer outcome for different Road Surface Conditions

The Fig. 4 depicts the outcome of Accelerometer for 5 road surface condition and also an idle case (Ignition is on but vehicle is not moving). It consist of three plots, first plot is for Acceleration (g) in X-axis, second is for Acceleration (g) in Y-axis, and next for Acceleration (g) in Z-axis. Fig 4 (a) is clearly depicting the vibration of the Idle vehicle and other road condition, in Fig. 4 (b) except smooth road and idle case in all other case the readings are high, in Fig 4 (c) maximum cases the readings are less than 10g because of gravitational force. From the CDF plot shown in Fig 8 we can identify that the magnetometer readings are different from each other for 5 road condition. This clear distinction of road surface condition is helpful when the classification algorithm classifies the categories of each road surface condition.

B. Performance Evaluation of Machine Learning Algorithms

The Accuracy and F1 Score manifest the performance of classifier. Accuracy is defined by correctly classified number of subject divided by total number of subjects in data set which is shown in Eq.1

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (2)$$

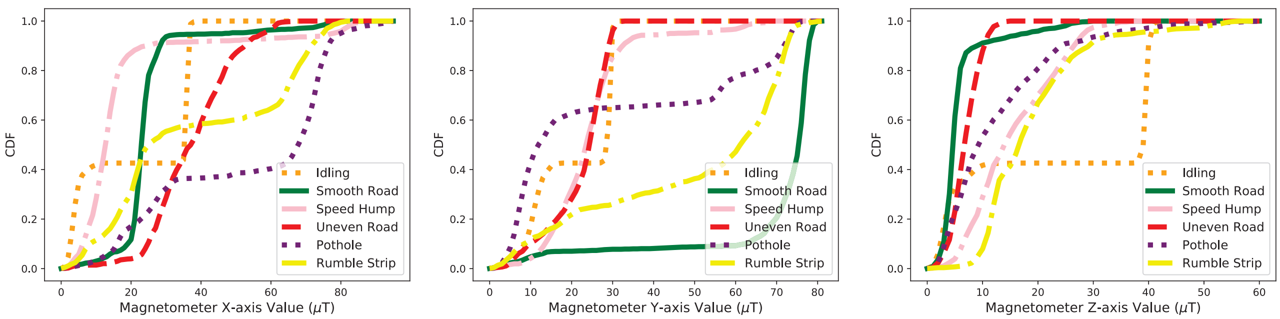


Fig. 8: CDF and MF values for various Road Conditions to validation Road condition

TABLE IV: Comparison between classifier and feature selection techniques Based on F1 Score

Feature Selection Techniques	Classifier		
	SVM	RandomForest	RandomTree
BestFirst	0.92	0.95	0.92
GreedyStepwise	0.91	0.94	0.91
Ranker	0.91	0.92	0.90
Average	0.91	0.94	0.91

Where TP - True Positive
TN - True Negative
FP - False Positive
FN - False Negative
We have measured the classification by F1 Score. F1 Score defines the harmonic average of Precision and Recall and it ranges from 0 to 1. Equation of F1 Score is shown in Eq.2.

$$F1Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{3}$$

The Classification result with extracted features using different machine learning model is shown in Fig 5 for SVM, Fig 6 for RandomTree and Fig 7 for RandomForest. The result in the form of accuracy and F1 Score is summarized in Table III and IV .

VII. Conclusion and Future Work

Smartphones have been determined efficient, reliable and affordable device when it comes to data collection and monitoring the road surface condition. To correctly classify the smartphone data, channeled into a machine learning algorithm for different road condition. For statistical analysis we have used three feature selection algorithm i.e Ranker algorithm, GreedyStepWise and BestFirst with, three classification algorithms i.e Support Vector Machine (SVM), RandomForest and RandomTree to access the performance analysis of the proposed model "MagTrack". "MagTrack" worked well in various distinguishable road condition with an average accuracy of "92%". In Future, we will develop a smartphone app for data Visualization which will provide an interface to display the road surface information to other travelers and road builders.

References

[1] L. Xiao and Z. Wang, "Internet of things: A new application for intelligent traffic monitoring system," *Journal of networks*, vol. 6, no. 6, p. 887, 2011.

[2] D. A. Johnson and M. M. Trivedi, "Driving style recognition using a smartphone as a sensor platform," in *2011 14th International IEEE Conference on Intelligent Transportation Systems (ITSC)*. IEEE, 2011, pp. 1609–1615.

[3] T. N. Schoepflin and D. J. Dailey, "Dynamic camera calibration of roadside traffic management cameras for vehicle speed estimation," *IEEE Transactions on Intelligent Transportation Systems*, vol. 4, no. 2, pp. 90–98, 2003.

[4] A. S. El-Wakeel, A. Osman, A. Noureldin, and H. S. Hassanein, "Road test experiments and statistical analysis for real-time monitoring of road surface conditions," in *GLOBECOM 2017-2017 IEEE Global Communications Conference*. IEEE, 2017, pp. 1–6.

[5] H. Sharma, S. Naik, A. Jain, R. K. Raman, R. K. Reddy, and R. B. Shet, "S-road assist: Road surface conditions and driving behavior analysis using smartphones," in *2015 International Conference on Connected Vehicles and Expo (ICCVE)*. IEEE, 2015, pp. 291–296.

[6] P. Mohan, V. N. Padmanabhan, and R. Ramjee, "Nericell: rich monitoring of road and traffic conditions using mobile smartphones," in *Proceedings of the 6th ACM conference on Embedded network sensor systems*. ACM, 2008, pp. 323–336.

[7] J. White, C. Thompson, H. Turner, B. Dougherty, and D. C. Schmidt, "Wreckwatch: Automatic traffic accident detection and notification with smartphones," *Mobile Networks and Applications*, vol. 16, no. 3, pp. 285–303, 2011.

[8] M. Collotta, G. Pau, V. M. Salerno, and G. Scatà, "A novel road monitoring approach using wireless sensor networks," in *2012 Sixth International Conference on Complex, Intelligent, and Software Intensive Systems*. IEEE, 2012, pp. 376–381.

[9] J. Liu, Y. Jia, and N. Xi, "Identification of road surface conditions based on laser scanning," in *2013 IEEE International Conference on Robotics and Biomimetics (ROBIO)*. IEEE, 2013, pp. 2476–2481.

[10] C. Zhang, L. Zhu, C. Xu, X. Du, and M. Guizani, "A privacy-preserving traffic monitoring scheme via vehicular crowdsourcing," *Sensors*, vol. 19, no. 6, p. 1274, 2019.

[11] A. Mohamed, M. M. M. Fouad, E. Elhariri, N. El-Bendary, H. M. Zawbaa, M. Tahoun, and A. E. Hassanien, "Roadmonitor: An intelligent road surface condition monitoring system," in *Intelligent Systems' 2014*. Springer, 2015, pp. 377–387.

[12] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg *et al.*, "Scikit-learn: Machine learning in python," *Journal of machine learning research*, vol. 12, no. Oct, pp. 2825–2830, 2011.

[13] D. Michie, D. J. Spiegelhalter, C. Taylor *et al.*, "Machine learning," *Neural and Statistical Classification*, vol. 13, 1994.

[14] J. Yu, H. Zhu, H. Han, Y. J. Chen, J. Yang, Y. Zhu, Z. Chen, G. Xue, and M. Li, "Senspeed: Sensing driving conditions to estimate vehicle speed in urban environments," *IEEE Transactions on Mobile Computing*, vol. 15, no. 1, pp. 202–216, 2016.