DESIGN AND IMPLEMENTATION OF PAVEMON:
A GIS WEB-BASED PAVEMENT MONITORING SYSTEM
BASED ON LARGE AMOUNTS OF HETEROGENEOUS
SENSORS DATA

A Thesis Presented

By

Salar Shahini Shamsabadi

to

The Department of Civil and Environmental Engineering

in partial fulfillment of the requirements

for the degree of

Master of Science

in the field of

Civil Engineering

Northeastern University
Boston, Massachusetts

December 2014
ACKNOWLEDGEMENTS

I am completely indebted to Professor Ralf Birken and Professor Ming Wang, for their mentorship, their guidance, and their valuable time spent generously with me. I am deeply grateful for their encouragement and support throughout my Master’s studies, which gave me the confidence that an international engineering student needs.

I would like to thank fellow researchers Dr. Yiyang Zhang, David Vines-Cavanaugh, Sindhu Ghanta, Yubo Zhao, Tarun Aleti Reddy, Yasamin Hashemi Tari, and all the other students in VOTERS research group, for their hard-work and their assistance in compiling some of the data presented in this thesis.

Lastly, I would like to acknowledge the National Institute of Standards and Technology (NIST) for funding the VOTERS project, allowing me and many others to receive financial support while working on exciting new technologies.
ABSTRACT

A web-based PAVEment MONitoring system, PAVEMON, is a GIS oriented platform for accommodating, representing, and leveraging data from a multi-modal mobile sensor system. Stated sensor system consists of acoustic, optical, electromagnetic, and GPS sensors and is capable of producing as much as 1 Terabyte of data per day. Multi-channel raw sensor data (microphone, accelerometer, tire pressure sensor, video) and processed results (road profile, crack density, international roughness index, micro texture depth, etc.) are outputs of this sensor system. By correlating the sensor measurements and positioning data collected in tight time synchronization, PAVEMON attaches a spatial component to all the datasets. These spatially indexed outputs are placed into an Oracle database which integrates seamlessly with PAVEMON’s web-based system.

The web-based system of PAVEMON consists of two major modules: 1) a GIS module for visualizing and spatial analysis of pavement condition information layers, and 2) a decision-support module for managing maintenance and repair (M&R) activities and predicting future budget needs. PAVEMON weaves together sensor data with third-party climate and traffic information from the National Oceanic and Atmospheric Administration (NOAA) and Long Term Pavement Performance (LTPP) databases for an organized data driven approach to conduct pavement management activities.

PAVEMON deals with heterogeneous and redundant observations by fusing them for jointly-derived higher-confidence results. A prominent example of the fusion algorithms developed within PAVEMON is a data fusion algorithm used for estimating the overall pavement conditions in terms of ASTM’s Pavement Condition Index (PCI). PAVEMON predicts PCI by undertaking a statistical fusion approach and selecting a subset of all the sensor measurements. Other fusion algorithms include noise-removal algorithms to remove false negatives in the sensor data in addition to fusion algorithms developed for identifying features on the road. PAVEMON offers an ideal research and
monitoring platform for rapid, intelligent and comprehensive evaluation of tomorrow’s transportation infrastructure based on up-to-date data from heterogeneous sensor systems.
# TABLE OF CONTENTS

ACKNOWLEDGEMENTS .................................................................................................................. ii
ABSTRACT .................................................................................................................................. iii
INTRODUCTION ............................................................................................................................... 1

An overview of Current Pavement Evaluation Methods .......................................................... 3
An Overview of Software Programs for Pavement Management Systems ............................... 9
Shortcomings of Pavement Evaluation Methods and PMS Software Programs ..................... 14

PART I: PAVEMON’S DESIGN ........................................................................................................... 17

VOTERS Pavement Condition Surveys ..................................................................................... 17
Statistical Data Fusion Approach for Pavement Condition Assessment .................................... 20
Temporal Data Fusion to Detect Noise in the Data ................................................................. 34
Supervised Machine Learning Approach for Feature Identification ....................................... 35
Decision Trees and Priority Matrices for Suggesting Maintenance Activities ...................... 39
Deterioration Model Considering Climate, Traffic, and Extreme Weather Events ............... 43
An Iterative Model for Predicting Future Budget Needs .......................................................... 54
Outlook ....................................................................................................................................... 56

PART II: PAVEMON’S IMPLEMENTATION ....................................................................................... 57

An Overview of VOTERS DATA .................................................................................................... 57
PAVEMON’s Architecture ............................................................................................................. 59
Georeferencing, Mapping, and Visualization ............................................................................. 61
Database Configuration ............................................................................................................... 67
Tools and Functions ...................................................................................................................... 68
PAVEMON’s Data Layers .............................................................................................................. 69
Examples of PAVEMON’s Functions .......................................................................................... 74
PaveMan Toolbox .......................................................................................................................... 79
Summary of PAVEMON’s Implementation Challenges .............................................................. 82

SUMMARY AND OUTLOOK ............................................................................................................. 85
REFERENCES .................................................................................................................................. 88
APPENDIX: PAVEMON’S ROUTING SYSTEM .................................................................................. 96
INTRODUCTION

Transportation infrastructures fuel the economic strength and social welfare by expediting everyday mobility of people, goods, and resources, consequently extending people’s business and social domains. They allow manufacturers to distribute their products to the appropriate markets quickly and inexpensively, enabling consumers to benefit from the lower prices of products and their higher qualities. Transportation facilities are also considered as a major influence on tourism development [1-3]. Therefore, it is not exaggerating to state that the road infrastructure is the backbone of the global economy and the key to the prosperity of the communities.

After World War II, roadway agencies focused on construction of the infrastructure networks to support the growing economy, the increasing popularity of automobiles, and the improved quality of life. Though pavements were designed to last long, they did not perform as expected; traffic loadings and environmental effects generate debonding, internal moisture damage and loss of subsurface support in pavements which accelerates their deterioration. Thus since the 1970s, agencies have shifted their focus from expansion to preservation to secure the socioeconomic roles of roadways.

United States’ Road Infrastructure Problem

With more than 4 million miles of roads and 600,000 bridges, US transportation infrastructure supports more than 3,000,000 million miles of travel each year [3, 4]. Even though the federal government spends around 50 billion dollars each year on roadways, ASCE’s evaluation in 2013 graded America’s road infrastructure only a D+. In fact, more than 1.3 million of US roads are in poor or mediocre conditions which is costing the economy a staggering amount of 101 billion dollars in wasted time and fuel each year. Additionally, ASCE report card indicates that approximately one-third of all the U.S. traffic fatalities are caused by poor pavement conditions. These statistics exhibit the monumental problem of infrastructure management in the scheduling and implementation of maintenance and repair operations in the US. As a matter of fact, this inefficiency in the
prioritization of maintenance expenditures within budgetary constraints is costing motorists 67 billion dollars each year [5].

**Pavement Management Systems**

Since the 1970s, highway administration has provided guidelines for developing a set of decision making tools known as pavement management systems. These guidelines were created to assist agencies in minimizing the overall costs of maintaining their transportation systems while maximizing the benefits of having a sound road infrastructure. Hudson et al. [6] describe a pavement management system (PMS) as “…a coordinated set of activities, all directed toward achieving the best value possible for the available public funds in providing and operating smooth, safe, and economical pavements.” PMS systems were developed based on the concept that resources for road maintenance can be efficiently managed by determining the current condition of the pavements and their deterioration rates. In fact, a successful PMS deploys a combination of engineering algorithms and business trade-off analysis techniques to suggest the right repairs, in the right place, and at the right time. Figure 1 shows the impact of timely maintenance on the money spent and life extension of roads [7, 8].

![Figure 1. Maintenance activities effect on lifecycle of pavements and money spent.](image)

**Components of a Pavement Management System**

Generally, pavement management systems consist of the following parts.

- A Comprehensive Database which stores the data required for the PMS analyses;
- Mathematical methods which develop useful decision making products;
- Calibration processes which fine-tune system’s parameters by using the latest field observations [9, 10].

A PMS process starts with collecting the data required for the PMS’s analyses and assigns Maintenance & Repair (M&R) activities with respect to the life-cycle costing and priorities as compared to other pavement requirements in the network. Every time new data is collected, this process reproduces the cost and condition predictions to recalibrate the system’s prediction models. Figure 2 summarizes a PMS process [11, 12].

Figure 2. Implementation process of a PMS.

Two key differences in PMSs are their method of evaluating pavement conditions and the software program developed to present the results and run the PMS models. Herein, first an overview of the state of the art pavement condition surveys is provided. Then, a review of prominent PMS software is presented.

**An overview of Current Pavement Evaluation Methods**

Collecting pavement condition data is typically a time-consuming and expensive process. Choosing the right pavement evaluation method is an integral part of any PMS. While M&R models has been steadily improving due to the advent of new referencing and visualization tools in addition to significant boost of computing power, progress made in
enhancing the data collection of a PMS in an affordable manner has not been as electric. Since the 1970s, PMSs are often constrained by outdated and low resolution condition data, forcing them to make lots of assumptions and extrapolations. Current methods of pavement inspections can be classified into two broad categories:

- **Manual Inspection**, which is “pavement condition data collection through processes where people are directly involved in the observation or measurement of pavement properties” [13]. Distresses assessments are made by either walking on the pavement (on-foot surveys) or while riding in a moving vehicle (windshield surveys).

- **Automated Mobile Inspection**, which is the “process of collecting pavement condition data by the use of imaging technologies or by other sensor equipment” [13]. Examples include profiling devices or 3D imaging techniques.

A review of some of the most common pavement evaluation methods is followed. These methodologies are:

- Pavement Condition Index (PCI)
- Pavement Surface Evaluation and Rating System (PASER)
- Condition Rating Survey (CRS)
- Localized Pavement Condition Metrics

**Pavement Condition Index (PCI)**

A widely-used method of road assessment in the United States is to calculate ASTM’s Pavement Condition Index (PCI). Developed in 1976 by the US Army of Corps, PCI is a numerical ASTM standard between 0-100 to indicate the general condition of pavements, with 100 being the best possible condition and 0 being the worst possible condition [15]. PCI inspections are either on-foot surveys (experts walk on each road and make measurements) or wind-shield surveys (experts ride in a car that is moving slowly and see through the windshield to record pavement distresses).

PCI methodology is well-documented in ASTM D6433, *Standard Test Method for Roads and Parking Lots Pavement Condition Index Surveys* [14]. As described in this
document, to calculate PCI experts visually inspect and assess the severity of 39 pavement distresses in three severity levels of low, medium or high. From these distresses, twenty are specific to flexible pavements, and 19 are specific to concrete pavements. PCI is then calculated in respect to severity of each of those reported distresses by using the measurements of the experts to deduct points out of 100. To calculate PCI experts should first divide the road network to three categories: individual roads or “branches”, pieces with similar work history or “sections”, and “samples”. Consequently, the pavement distresses stated earlier are inspected in these samples and their severity levels are determined. In flexible pavements, which are the focus of this thesis, the samples’ sizes are 2,500 ft² ± 1,000.

After the existing distresses and their severities are determined in the samples, the numbers are entered into a sheet from which the deduct values stated earlier are calculated. These calculations are based on available graphs in the ASTM specifications. Consequently, the total deduct values are subtracted from 100 and a weighted PCI average of inspected samples are calculated as the pavement condition index of each section. Figure 3 shows a graphic illustration of PCI methodology.

PCI scores are typically translated into the maintenance categories shown in Table 1 [13, 15].
Table 1. PCI Ratings Related to Maintenance and Repair Strategies.

<table>
<thead>
<tr>
<th>PCI</th>
<th>Maintenance Action Required</th>
</tr>
</thead>
<tbody>
<tr>
<td>90 &lt; PCI</td>
<td>Do Nothing</td>
</tr>
<tr>
<td>70 &lt; PCI &lt; 90</td>
<td>Preventive Repair</td>
</tr>
<tr>
<td>40 &lt; PCI &lt; 70</td>
<td>Rehabilitation</td>
</tr>
<tr>
<td>PCI &lt; 40</td>
<td>Reconstruction</td>
</tr>
</tbody>
</table>

Pavement Surface Evaluation and Rating System (PASER)

The Pavement Surface Evaluation and Rating System (PASER) is another popular measure of pavement conditions. PASER is a classification metric between 1-10 which indicates the general condition of pavements, with 10 being the best possible condition and 0 being the worst possible condition [16]. The PASER rating system uses four major surface distresses based on visual inspections of the experts on the field:

- Surface defects (Raveling, flushing, polishing)
- Surface deformation (Rutting, distortion)
- Cracks (Transverse, Longitudinal, Alligator, Block, Reflection, Slippage)
- Patches & potholes

Surveyors first assess the general condition of the pavement to calculate PASER, i.e. if the pavement is in a generally good, fair or poor condition. Then, they evaluate the four pavement distresses mentioned above by referring to the snapshots of each rating, an example is shown in Figure 4 [16]. The PASER rating scale can generally be translated into the maintenance categories shown in Table 2. PASER is an estimated measure of road conditions which is subjective to the surveyor’s judgment.

Table 2. PASER Ratings Related to Maintenance and Repair Strategies.

<table>
<thead>
<tr>
<th>PASER</th>
<th>Maintenance Action Required</th>
</tr>
</thead>
<tbody>
<tr>
<td>PASER = 9, 10</td>
<td>Do Nothing</td>
</tr>
<tr>
<td>PASER = 8</td>
<td>Little Maintenance</td>
</tr>
<tr>
<td>PASER = 7</td>
<td>Routine Maintenance</td>
</tr>
<tr>
<td>PASER = 5, 6</td>
<td>Preventive Repair</td>
</tr>
<tr>
<td>PASER = 3, 4</td>
<td>Rehabilitation</td>
</tr>
<tr>
<td>PASER = 1, 2</td>
<td>Reconstruction</td>
</tr>
</tbody>
</table>
Condition Rating Survey (CRS)

The Condition Rating Survey (CRS) is another popular metric to rate the overall condition of pavements. CRS is a numerical metric between 1-9 to indicate the general condition of pavements, with 9 being the best possible condition and 1 being the worst possible condition [17]. The CRS rating scale can generally be translated into the maintenance categories shown in table 3.
Table 3. CRS Ratings Related to Maintenance and Repair Strategies.

<table>
<thead>
<tr>
<th>CRS</th>
<th>Maintenance Action Required</th>
</tr>
</thead>
<tbody>
<tr>
<td>7.6 ≤ CRS ≤ 9.0</td>
<td>Do Nothing</td>
</tr>
<tr>
<td>6.1 ≤ CRS ≤ 7.5</td>
<td>Preventive Repair</td>
</tr>
<tr>
<td>4.6 ≤ CRS ≤ 6.0</td>
<td>Rehabilitation</td>
</tr>
<tr>
<td>1.0 ≤ CRS ≤ 4.5</td>
<td>Reconstruction</td>
</tr>
</tbody>
</table>

CRS values differ in 0.1 increments, an example is shown in Figure 5 [18] where a road with CRS of 5.9 has been compared to a road with CRS of 5.8. While the surveyors may use the images provided as examples of pavement conditions in terms of CRS, CRS can also be derived from several algorithms which relate other measurements such as Roughness and Rutting to CRS. These algorithms, however, are generally dependent on the local agency which developed them for its own purposes.

Figure 5. Comparison of CRS = 5.9 with CRS = 5.8 [18].

Localized Pavement Condition Metrics

In addition to the well-documented methods outlined above, many agencies have created customized rating systems that specifically meets their needs. Some of these rating systems simply classify the roads in three categories of good, fair, and poor, while the others
are more sophisticated [18]. Major detriments of these localized systems are that they are difficult to compare the ratings of other types of systems, and it is often hard to convert these ratings to others as there were not much research done based on those metrics.

**An Overview of Software Programs for Pavement Management Systems**

User-friendly software that is capable of presenting pavement evaluation results and carrying the necessary processes is an integral part of a successful Pavement Management System. Hence some of the most popular PMS software programs currently used by the roadway agencies were reviewed. These programs are:

- MicroPAVER
- StreetSaver
- PAVEMENTview
- RoadSoft GIS

**MicroPAVER**

Developed by the US army of corps in the 1980s, MicroPAVER is probably the most common software used for PCI-based systems. MicroPAVER stores road network inventories from which it calculates condition ratings. Micropaver also has several decision making tools for scheduling M&R operations in a cost-effective manner. These tools create present and future pavement condition predictions, and suggest maintenance activities accordingly. Additionally, Micropaver allow the user to analyze different budget scenarios and customize some of the assumptions made by the software [19].

PCI distresses and their corresponding severity levels are the main inputs of Micropaver, from which the program calculates the PCI. Comments and snapshots of the pavements can be manually entered for the surveyed sections. Micropaver groups roads with similar traffic and climate conditions into “families”, and as shown in Figure 6, assumes the same deterioration rate for each family.
Figure 6. MicroPAVER family-based prediction model (left), condition analysis outputs displayed in MicroPAVER’s GIS interface (right).

Micropaver has a Geographic Information System (GIS) module which allows visualizing the road conditions in maps, tables and graphs, like those shown in Figure 6 [18]. MicroPAVER provides four tools to help in managing M&R operations. These tools are:

- Summary charts: Summaries of pavement information layers in graphs for different distresses
- Standard reports: List of pavement branches in addition to their repair history and current conditions.
- Re-inspection reports: Reports after new inspection data has been entered into the software.
- GIS reports: Maps, tables, and graphs of future condition predictions and budget needs [19].

**StreetSaver**

Developed by the Metropolitan Transportation Commission (MTC), StreetSaver is used by more than 350 agencies in the US alone [20]. StreetSaver, like Micropaver, is based on PCI methodology. The PCI used by this software, however, is calculated from seven short-listed distress types instead of the thirty-nine pavement distresses used in Micropaver. These distresses are based a study done by Shahin et. Al. [21] who has summarized the original PCI distresses in seven categories. For flexible pavements, these distresses are longitudinal & transversal cracking, block cracking, alligator cracking, distortions, rutting &
depressions, patching, and weathering. As shown in Figure 7 [22], StreetSaver creates visualization layers for historical or current PCI data.

![Figure 7. StreetSaver Pavement Condition GIS Report [22].](image)

StreetSaver suggests the required repairs based on PCI. This software allows the user to create different budget scenarios and compare the benefits of each by storing several scenarios in the same session. Furthermore, some of the assumptions made by StreetSaver for suggesting M&R operations can be customized by the user. StreetSaver can create over thirty different graphs and reports of pavement conditions and future budget needs.

**PAVEMENTview**

PAVEMENTview is part of the Cartegraph software which was developed for asset management (signposts, utilities, etc.). This software inventories the pavement networks’ conditions and geometry details (length, width, etc.) and allows to create Capital improvement planning (CIP) scenarios, i.e. budget allocation plans for bringing up the pavement condition of road network after a user-defined number of years. PAVEMENTview suggests repair strategies and allows the user to customize its decision trees. This software also calculates the priorities of repairs based on benefit-to-cost ratio of each maintenance activity.
PAVEMENTview contains a comprehensive help document for assisting the surveyors in distress inspections. PAVEMENTview calculates a metric named the Overall Condition Index (OCI), a similar metric to PCI, after the user enters the observed pavement distresses. In addition to predicting the future condition of the network, PAVEMENTview estimates the remaining life of the pavement sections. Snapshots of the user-interface of PAVEMENTview and its graphical GIS report is shown in Figure 8.

![Figure 8. PAVEMENTview user and GIS interfaces [24].](image)

**RoadSoft GIS**

Developed by the Michigan Tech Transportation Institute, Roadsoft GIS is a PASER-based software heavily used in the cities and villages of Michigan. In addition to pavement information layers, RoadSoft stores asset management information such as signposts and utilities. As Roadsoft is compatible with laptops and tablets to calculate PASER from distress sheets, it is also used in data collection [25].

Figure 9 [26] shows Roadsofts’ deterioration curve to a family of roads in a sample map. The graph shows Road Soft’s prediction for the remaining service life of the pavements.
RoadSoft allows the user to input additional data such as traffic counts or traffic crashes data and incorporate them in the reports generated by the system. Figure 10 [27] shows such interfaces of Roadsoft program. Roadsoft’s reports can be exported in standard GIS formats to be used in other programs.
Shortcomings of Pavement Evaluation Methods and PMS Software Programs

State-of-the-art road inspection methodologies currently practiced by the cities generally have the following shortcomings:

- Time-consuming in performing the surveys and processing the data.
- Expensive to be carried out on a frequent basis.
- Subjective to human errors.
- Inconsistent, i.e. if two experts are to rate the same pavement at the same time they might have different assessments.
- Mostly require traffic blockage, makes these surveys costly (both to the agencies and road users) to perform especially in dense areas.

Consequently, as these surveys are performed years apart they quickly become outdated, causing the decisions derived from them to be less valuable. Implementation of a pavement management system that can make the proper assessment of maintenance operations and priorities needs to be armed with an efficient road inspection method that is non-intrusive (e.g. does not require traffic blockage), is fast, and requires minimum manual effort so that it can be performed regularly [28].

In addition to the road inspection methodologies, the state of the practice software programs have the following limitations:

- Installation required. Data, report, graphs and other features cannot be accessed without the necessary software installations.
- Experts required for updating the data. Each time new data is collected people familiar with the software need to manually update the database.
- Subject to human errors. As the software is driven by human, mistakes are inevitably made from time to time.
- Limited Visuals. This is because of the limited data provided by the surveys. Powerful visuals can make a dramatic effect on the decisions made by seeing the consequences of each action.
• No Verification Method. After the data has been placed into the database, there’s not a way to verify the data.

• Expert-opinion Decision Method. Decision Making algorithms are mostly derived from intuition and experience rather than data. Consequently, one decision that might work for one agency might not work for another.

PAVement MONitoring system, PAVEMON, armed with Versatile Onboard Traffic Embedded (VOTERS) inspection method has been developed to overcome these shortcomings. PAVEMON’s main advantages are its:

• data driven nature. Rather than intuition or personal experience, PAVEMON’s mathematical models are all compelled by immense amounts of information made available for every meter of every inspected road.

• zero-installation. Users don’t need to do any software installations to operate PAVEMON.

• easy access to record of road’s condition from anywhere via the internet.

• automated information update. After each survey, data are automatically processed, placed into the database and updated on PAVEMON.

• verification capabilities. For each meter of the road, images of the pavement are geo-tagged and placed on map to provide high resolution ground truth information.

• powerful visuals, makes for dramatic presentation of road conditions and the ability to challenge the M&R decisions made effectively.

• customizable nature, system can be customized based on agencies’ preferences and goals.

• high frequency of data update, analyses models are calibrated in short time intervals.
The remainder of this thesis is structured in two parts: “PAVEMON’s Design” and “PAVEMON’s Implementation”.

In the first part (PAVEMON’s Design), the focus is on the methodologies and algorithms used by PAVEMON. In this chapter, the VOTERS inspection methodology is introduced. Then, PAVEMON’s statistical data fusion model for assessing PCI is elaborated. Other fusion algorithms created within PAVEMON for noise-removal and feature identification are also discussed. Then, the data driven decision making methodologies developed for leveraging pavement information layers to plan optimum M&R operations are explained.

In the second part (PAVEMON’s Implementation), the methodologies introduced and validated in the first chapter are implemented. Consequently, the developments made to georeference VOTERS data and weave them with third-party climate and traffic information layers are discussed. The structures of PAVEMON’s different components are explained and features created on PAVEMON are listed. Functionalities of PAVEMON are shown through some examples at the end.
PART I: PAVEMON’S DESIGN

Due to the shortcomings of the current pavement inspection methodologies stated earlier, these surveys are typically performed years apart and quickly become outdated, causing the decisions derived from them to be less valuable. Implementation of a pavement management system that can make the proper assessment of maintenance operations and priorities needs to be armed with an efficient road inspection method that is non-intrusive (e.g. does not require traffic blockage), is fast, and requires minimum manual effort so that it can be performed regularly. Such inspection method has been developed by the Versatile Onboard Traffic Embedded (VOTERS) project.

VOTERS Pavement Condition Surveys

The VOTERS project has developed a multi-modal mobile sensor system [82] capable of collecting pavement-related information which includes

- Acoustic technology that uses tire-induced vibrations and sound waves to determine surface texture, roughness and overall condition. The waves are recorded with directional microphones [29] and a newly developed Dynamic Tire Pressure Sensor (DTPS) [30, 31].
- Millimeter-wave radar technology for the near-surface inspection of roadways and bridge decks focusing on determining road profile and rutting depth in addition to detecting surface defects and features such as potholes, water, or metal (manholes and utility covers) [32].
- Video technology used to capture surface defects for verification of results from other sensors and analyzing cracks types and intensity [33, 34].
- Improved air-coupled Ground Penetrating Radar (GPR) array technology that maps subsurface information such as the pavement layers (thicknesses and electromagnetic properties) in addition to rebar corrosion and delamination of bridge decks [35-37].
A prototype vehicle has been outfitted with VOTERS sensor system (Figure 11) to monitor road conditions as it is roaming through daily traffic. These data are processed to render meaningful knowledge of the road, i.e. parameters such as road profile, crack density, road roughness, and micro texture depth.

![Overview of the VOTERS survey vehicle and sensors.](image)

Figure 11. Overview of the VOTERS survey vehicle and sensors.

Following parameters are the main measures calculated from the VOTERS system prior to data being used by PAVEMON.

*Mean Texture Depth*

International Organization for Standardization (ISO) describes pavement texture as “the deviation of a pavement surface from a true planar surface” [38]. MTD shows severity of segregation and raveling, two dominant types of pavement distresses. VOTERS estimates MTD from its microphone signals which is thoroughly explained in [29].
*International Roughness Index*

ASTM defines IRI as “a quantitative estimate of a pavement property defined as roughness using longitudinal profile measures” [39]. VOTERS estimates IRI from its dynamic tire pressure signals as explained in [40].

*Crack Density*

Cracks are one of the prominent pavement distresses. Images of VOTERS camera are corrected for distortion due to the angle of projection of the camera and then analyzed for cracks. A Hessian-based multi-scale filter has been applied to detect ridges in the image at different scales. Crack density is the main information rendered after processing these images [34].

*Rutting Depth*

Pavement rut depth has a direct influence on traffic safety in addition to water run-off issues which generally appears due to the weak load bearing capacity of the pavement. VOTERS can estimate rutting depth by detecting changes in the measurements of its 5 channel array of mm-wave radars placed underneath the vehicle [32].

*Sub-layer’s Thickness and Material*

Loss of subsurface support and internal moisture is one of the main reasons for pavement’s aging. VOTERS air coupled GPR system provides subsurface defects such as corroded rebar, trapped moisture, voids, and the pavement layers thicknesses at traffic speed [35-37].

Table 4 summarizes these sensors and their measurements.
### Table 4. VOTERS Sensors and Measurements.

<table>
<thead>
<tr>
<th>Technology</th>
<th>Measurements</th>
<th>Specifications</th>
</tr>
</thead>
</table>
| Directional Microphone              | • Mean Texture Depth (MTD)                        | Sensor height: ½ - 3 inch  
Sampling Rate: 2 - 200 KHz  
Sensitivity: 44 - 52 mv/Pa                                                   |
| Dynamic Tire Pressure Sensor (DTPS) | • International Roughness Index (IRI)  
• Road ProfiRoad Height Variations | Frequency: 0.5 Hz - 20 KHz  
Sampling Rate: 2 - 200 KHz  
Dynamic Pressure: 0 - 1 psi                                                   |
| Camera                              | • Crack Type (Alligator, Transversal, Longitudinal)  
• Crack Density                 | Resolution: 2.82 Megapixel  
Speed: Gigabit Ethernet                                                   |
| Millimeter-Wave Radar               | • Rutting depth                                    | Operation: 24 GHz  
Arrays: 5 channels                                                        |
| Ground Penetrating Radar            | • Sub-layers’ thickness  
• Sub-layers’ materials  
• Subsurface Feature Identification (delamination, potholes, etc.)  
• Subsurface Moisture | Frequency: 0.8 - 5 GHz  
Data rate: 1000 trace/sec  
Low cost  
Low power  
Small                                             |

**Statistical Data Fusion Approach for Pavement Condition Assessment**

While each of the parameters described before provides worthy knowledge about an aspect of the road condition, none alone can furnish sufficient information that would encompass all aspects of pavement condition. PAVEMON, like other PMS systems described earlier, requires a performance metric to be able to derive inferences and decisions for the front end user. As VOTERS is aiming to survey the cities, one of the prominent pavement indicators discussed earlier should be chosen. Since PCI is the most widely-used parameter in the Northeast, this measure is selected for assessment of overall pavement conditions. Purpose here is to predict PCI from a combination of the VOTERS sensors’ measurements. Consequently, instead of having many metrics for each road, all information will be summarized in one, as illustrated in Figure 12. Red, yellow, and green colors in this figure indicate the parameter is not in the desirable range for a good pavement.
A methodology that can successfully deal with heterogeneous and redundant observations is to fuse them for jointly-derived higher-confidence results. This process is often referred to as “data fusion” or “sensor fusion” or “information integration”. The US Department of Defense defines data fusion as “a multilevel, multifaceted process dealing with the automatic detection, association, correlation, estimation, and combination of data and information from multiple sources “. There is a need for data fusion since often times, one sensor is not capable of providing reliable information over all arrays and environments. While data fusion has been used widely in different applications, what runs through all of them is that their purpose is achieving a better inference than what each data source alone could achieve.

The concept of multi-sensor data fusion is hardly new. Along with the emergence of new sensors and innovative processing algorithms, multi-sensor data fusion has been extensively applied in military, medical, engineering and other applications. Some possible objectives of data fusion are recognizing an environmental condition, detecting an object presence, navigating an autonomous vehicle, collecting immediate information after a disaster and monitoring an infrastructure status [41- 44].

A naive way of combining information is by concatenating the features as a single input vector to a standard classifier. This is suboptimal because feature types may be incompatible and information structure can be lost. One might assume that combining data
from heterogeneous sources will always offer more reliable information. This, however, is seldom the case. Output from a multi-sensor data fusion system might have less accuracy than what the most appropriate sensor in the sensor suite could furnish. This could be as a result of having contradictory information or presence of noise in datasets. Another possible reason could be the existence of redundant information, giving more weight to one facet of knowledge as several parameters could be representing the same variable in different ways [43, 44].

Data fusion algorithms have been developed to deal with these challenges. Pattern recognition, artificial intelligence, signal processing and regression are some common fusion techniques. Deploying the most appropriate algorithm is decided from what is expected from the fusion system. Here, an algorithm which is adept at variable selection could be suitable as it is likely to be concluded that not all of these sensors are necessary to predict the PCI scores. Furthermore, as PCI is a physical measurement, the final model is preferred to be intelligible where it can be physically justified why it works.

Machine learning techniques such as Neural Networks and Support Vector Machine were considered first. While having excellent training capabilities and dealing well with outliers, these methods generally have the shortcoming of complexity. Consequently, while the outputs from these algorithms might work, it might be difficult to explain the rationale behind it. Moreover, they often depend more heavily on training than other potential methods, specifically regression-based techniques [45, 46].

**Proposed Data Fusion Model**

Montgomery et al. [47] defines regression analysis as “a collection of statistical tools that are used to model and explore relationships between variables that are related in a nondeterministic manner”. Strengths of these models are that they often do not require as much training as most of the machine learning techniques do, and the results could possibly be justified through the meaning of the parameters of the final model. The simplest case of regression analysis includes one scalar predictor variable and one scalar response variable. Having multiple scalar or vector predictors is known as multiple regressions. A majority of regression models involve multiple predictors [45, 47].
A well-suited fusion model for the application here should be able to identify a smaller subset of original predictors which can be the main contributors in PCI prediction when included jointly in a model. As Multiple Regression does not have this functionality, other regression-based techniques that are able to penalize the size of the regression coefficients were considered, algorithms such as Ridge Regression (penalizing through least squares), Best Subset Regression (penalizing through statistical tests such as correlation coefficient) and Stepwise Regression. Strength of these methods is most striking in the presence of multi-collinearity as they have a variable selection process.

The fusion model used here is stepwise regression. As it could be inferred from the name, Stepwise Regression approaches the problem by entering and removing variables one at a time. Consequently, a predictor stays in the model only if it boosts the model’s fit to the data significantly. One approach to define this significance is to use an F-test, but other techniques such as t-tests, adjusted R-square, Bayesian information criterion, or false discovery rate are also possible [45, 47, 48].

Properly used, Stepwise Regression can potentially select a small subset of independent variables with a high prediction capability among many by poking variables in or out. Typically, if some variables have many fewer observations than the others, Stepwise Regression may converge on a poor model as these variables will decrease the test period for all the models [48, 49]. Here, this is not the case and Stepwise Regression was selected for fusing VOTERS parameters to estimate the PCI.

**Predictors**

In addition to Mean Texture Depth (MTD) and International Roughness Index (IRI) discussed earlier, Standard deviation and Average Energy (area under the frequency domain curve over a certain band) of raw data for DTPS, Accelerometer, and Laser Height Sensor were included in the Stepwise Regression model as a possible predictors. From intuition and observations, disturbance in the pavement conditions would cause variations in the raw signal of these sensors, hence they could be potential candidates for predicting PCI. Example of how the signal of DTPS raw data hence the Standard Deviation and
Average Energy values change with pavement conditions are shown in Figure 13. Similar observations were made for the Accelerometer and Laser Height sensors.

![Figure 13. Raw data with processed Standard Deviation and Average Energy of DTPS sensor for three roads with different pavement conditions.](image)

Additionally, Zhang et al. [84] indicated that pavement conditions could be estimated through standard deviation of the amplitudes of frequency spectra from the acoustic measurements. Hence this parameter has also been calculated and used as a potential PCI predictor. In total ten parameters were entered into the fusion model. Figure 14 shows a graphic presentation of the proposed fusion model with all the sensors and the predictors derived from them.
In the first step, only one predictor is used to estimate the response (PCI). This predictor is the one with the highest correlation with the PCI. With the addition of a new variable at each step, a new multiple regression model to the existing variables will be fitted. Then it will be determined whether predictor variables have gained a statistically considerable improvement on predicting the response variable with this new addition. This is done by testing the hypotheses $H_0^j: \beta_j = 0$ vs. $\beta_j \neq 0$ for the new jth variable. If $H_0^j: \beta_j = 0$ is not rejected, i.e. if $\beta_j$ is not significantly different from zero, this indicates the corresponding variable $x_j$ does not increase the prediction capability of the model and it may be dropped. A confidence interval (CI) for each $\beta_j$ in the final model can be calculated to assess the effect of each $x_i$ on the response variable. The $\beta_i$ for the other variables are also calculated when a new variable enters the equation, as this new addition might have caused another variable to turn into an insignificant contributor.

Partial $F_i$ cited above is computed as:

$$r^2_{y|x_p|x_1,\ldots,x_{p-1}} = \frac{SSE_{p-1} - SSE_p}{SSE_{p-1}} = \frac{SSE(x_1,\ldots,x_{p-1}) - SSE(x_1,\ldots,x_p)}{SSE(x_1,\ldots,x_{p-1})}$$
Where:

\( r \) : Total variation in \( y \) that is accounted for by a regression on \( x \)

\( \text{SSE}_p \) : Error sum of squares in step \( p \) (on variables \( x_1, \ldots, x_p \))

Supervised fusion models need training and reference PCI values for training the
data were furnished by CDM Smith for the Brockton, MA.

**Experiment**

In August 2012 the VOTERS prototype van (Figure 11) drove on a route in
Brockton which was planned to contain various ranges of PCI values from 0 to 100, a
decent test bed to provide the fusion model with the boundaries it needs. Surveyed roads
were sixteen two-lane streets. The prototype vehicle drove on these roads three times in
each direction, i.e. each lane three times. Each of the parameters explained before were
calculated for each run on these streets, then each two runs on different lanes of a street
were paired and the values were averaged for both lanes. At the end, all these numbers
were normalized and a total of 48 observations were available (three per street).

**Data Association**

As described, ten parameters were used as inputs to the fusion model. Here, the
model iterates over different permutations of these using Stepwise Regression and renders
a subset of them based on the partial F-tests calculated in each step. Partial F-Tests in each
step are calculated with the model’s variables and the reference PCI values. From the
sixteen streets surveyed, eight were entered into the algorithm with all the ten parameters
calculated for these streets to train the fusion model. This is a total of 24 observations
(vehicle has driven three times in each direction). These streets and their reference PCI
scores are shown in Table 5. These specific streets were chosen for training in order to
entail all ranges of PCI to fine-tune the fusion model’s boundaries.
Table 5. PCI of the streets used to train the fusion algorithm.

<table>
<thead>
<tr>
<th>TRAIN SET</th>
<th>Randolph Rd</th>
<th>Boyle Rd</th>
<th>Boundary St</th>
<th>Christopher St</th>
<th>Ridge St</th>
<th>Intervale St</th>
<th>Swtell Ave</th>
<th>Hovdn Ave</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reference PCI</td>
<td>0</td>
<td>10</td>
<td>14</td>
<td>31</td>
<td>55</td>
<td>75</td>
<td>75</td>
<td>90</td>
</tr>
</tbody>
</table>

In the final model, equation 2, three parameters were selected by the fusion model to predict PCI values, i.e. the presence of any other parameter with the existence of these three was insignificant or harmful to the correlation rendered by the model. These parameters come from the two acoustic sensors of Microphone and DTPS. The parameters, as shown in equation below, are Mean Texture Depth, Standard Deviation of FFT of Microphone, and Standard Deviation of DTPS Raw Data. Later, an attempt will be made to justify the results through physical meanings of these parameters.

\[
\text{PCI} = 46.5 - 1.9N_{\text{StdFFT}} - 9.28N_{\text{MTD}} - 5N_{\text{StdDTPS}}
\]

Where:

\[
N_{\text{StdFFT}} = \text{Normalized Standard Deviation of FFT of Microphone}
\]

\[
N_{\text{MTD}} = \text{Normalized Mean Texture Depth}
\]

\[
N_{\text{StdDTPS}} = \text{Normalized Standard Deviation of DTPS Raw Data}
\]

PCI values of the remaining 24 sections (8 streets) were predicted with the fusion model developed above. Results are provided in Figure 15. It should be noted that the results were very consistent in the three runs and the numbers in Figure 15 are the averaged values of the three runs.
According to Figure 15, considerable discrepancy exists between the predicted PCI and reference PCI values of street #7. Looking at the images collected from the camera behind the VOTERS van (Figure 11), it could be seen that the real condition is in fact closer to what was predicted by the proposed fusion model.

The reference PCI was from a city-wide manual survey performed in 2006. Using a deterioration model (Micropaver’s Markov chain deterioration model, [19]) these values were forward projected for the year the VOTERS survey was conducted in 2012. It seems that street #7 deteriorated more severely than expected which could be an impact of severe weather conditions such as snowstorms or floods [50]. Other differences seem negligible and is difficult to observe which number is closer to the real condition with the images. Setting aside Street #7, fusion model results have more than 97% correlation with the reference PCI scores, as shown in Figure 16.
Figure 16. Final Fusion Model and its Correlation with Reference PCI.

Physical Justification

PCI is a parameter calculated from the severity of twenty pavement distresses. Shahin et al. [21] specifies seven distresses for this calculation and indicates they entail all the nineteen original ones. Alligator Cracking, Block Cracking, Longitudinal and Latitudinal Cracking, Distortions, Weathering and Raveling, Patching, and Rutting and Depressions are these seven shortlisted distresses.

The purpose of this section is to justify the final fusion model through physical meanings of its coefficients. It is not claimed here that the presence of these distresses can be predicted through the sensors’ measurements but rather discussing how these seven distresses are present in the parameters of the proposed model.

1,2,3. Alligator Cracking, Block Cracking, Longitudinal and Latitudinal Cracking: Friction of the road differs when cracks are present. This would be reflected in MTD and possibly Std_FFT. In Figure 17, MTD values are compared in three street sections: a section with no cracks, a section with average longitudinal and latitudinal cracks and a section with severe alligator cracks.
4. **Distortions:** Distortions cause variations in road profile causing StdDTPS to generate different values. An example is shown in Figure 18.

![Figure 17. Impact of Cracks on MTD values](image)

5. **Weathering and Raveling:** Weathering and raveling are the wearing away of the pavement surface typically as a result of displaced aggregate particles and are indicators of hardened binder or a poor mixture quality. MTD could reflect this distress as it is an indicator of the microtexture of the pavement. An example was shown in Figure 18.

![Figure 18. Raw and normalized measurements of (Left) mic after driving over a course and fine pavement, (Right) DTPS after driving over rough and smooth pavements](image)

6. **Patching:** Patches also affect road friction hence the microphone data. It also might be reflected in Std_DTPS as it changes elevation of the road and the pressure in the tire. Two 10-meter segments of the same street have been compared in presence and absence of
patches in Table 6 and a significant drop is observed in Std_DTPS when patching is present.

Table 6. Comparison of Std_DTPS values in presence and absence of patches of a street

<table>
<thead>
<tr>
<th>Average Std_DTPS for 10 meters</th>
<th>0.01432108</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average Std_DTPS for 10 meters</td>
<td>0.02227892</td>
</tr>
</tbody>
</table>

7. Rutting and Depressions: Rutting and Depression happen when the pavement surface areas with slightly lower elevations than the surrounding. StdDTPS might entail this distress, as rutting and depressions are changes in road profile and might affect the pressure inside the tire. In the survey conducted, no significant rutting was present in the surveyed streets thus there were no data to substantiate the statement here.

Validation and Accuracy: Second Brockton Field Test

In November 2013, December 2013 and January 2014 VOTERS surveyed two hundred (200) miles of road in Brockton City, MA, including fifteen of the streets surveyed in August 2012. The route and PCI results from the fusion model are shown on the right side of Figure 19. Using the fusion model elaborated in this paper, PCI values have been calculated from equation 2. Having this new dataset, the accuracy of the proposed model was studied with three methods:

1. Comparing to the camera images;
2. Comparing to the manually collected PCI scores, forward projected from 2006 for validation and accuracy;
3. Comparing to PCI scores of the fifteen streets surveyed in both 2012 and 2013-14 to check the consistency and repeatability.
Accuracy Check 1: Comparison with Camera Images

A simple way to roughly validate system accuracy is to check if the VOTERS camera corresponds to the conditions determined using the PCI fusion model. This type of manual validation was performed for the entire Brockton survey. The agreement was strong; see examples on Figure 19.

![Figure 19. PCI Map of Second Brockton Test (Left), Validation with Images from Camera (Right).](image)

Accuracy Check 2: Comparison to an Existing Condition Survey

A quantitative way to validate system accuracy is to compare to an existing professionally done condition survey, such as the one VOTERS has access to from CDM Smith. PCI scores of CDM Smith were forward projected from 2006 to the year of VOTERS survey using Micropaver, and then compared to the VOTERS PCI scores. Findings from this comparison are shown in Table 7.
Table 7. Comparison of predicted PCI in second Brockton field test to the reference PCI.

<table>
<thead>
<tr>
<th></th>
<th>Predicted PCI</th>
<th>Reference PCI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average Condition</td>
<td>60</td>
<td>56</td>
</tr>
<tr>
<td>Percent &gt; critical condition (PCI = 55)</td>
<td>67%</td>
<td>62%</td>
</tr>
<tr>
<td>Percent ≤ critical condition (PCI = 55)</td>
<td>33%</td>
<td>38%</td>
</tr>
<tr>
<td>Correlation</td>
<td></td>
<td>0.71</td>
</tr>
</tbody>
</table>

Comparisons made in Table 7 in addition to the high correlation of 0.71 between the two datasets suggest that the predicted PCI scores are reasonably accurate. This correlation is not perfect as the CDM Smith values were forward projected from 2006 thus not the current exact condition.

**Accuracy Check 3: Comparison of two Consecutive Years**

A comparison analysis between fifteen streets which were repeated in both surveys was conducted to demonstrate that the proposed fusion model is consistent, and is sensitive to deterioration over time. Consistency can be shown by demonstrating that the overall trend of results from year one is close to year two. This was expected as there were no significant repairs and yearly deterioration should be relatively low. Figure 20 entails the results of this comparison study.

![Figure 20. Comparison of PCI predictions in two consecutive years.](image-url)
From the high correlation of 0.72 it can be inferred that the predicted PCI is consistent. Sensitivity to deterioration over time – evident from the negative sign of the average difference- is another conclusion inferred from the above analysis. Year two predicted PCI values suspiciously indicate that some roads have improved which is due to minor repairs done by the city since year one.

**Limitations**

After practicing PAVEMON’s PCI fusion algorithm, it was observed that it has the following limitations:

- Requires a minimum speed of 15 mph.
- Sudden change of speed (high acceleration) affects the predictions.
- Manholes could be misinterpreted as distress.

For the second and third limitations, noise-removal algorithms were developed within PAVEMON to enhance the prediction results.

**Temporal Data Fusion to Detect Noise in the Data**

As mentioned above, when VOTERS vehicles goes over a manhole, there’s a chance of false prediction by the fusion model. Manholes could be easily detected by VOTERS mm-wave radar, thus given the accuracy in timing, Figure 21 illustrates data fusion opportunities through temporal correlation. Acoustic data streams are fused with mm-wave reflectivity data to differentiate false negatives from true negatives.
Figure 21. Detecting false negatives in acoustic data through temporal data fusion

Consequently, these false negatives have been identified, the affected areas were removed and the PCI values were recreated by spatial averaging between two ends of the removed sections.

**Supervised Machine Learning Approach for Feature Identification**

In addition to overall ratings, features on the road need to be identified for a better pavement evaluation. Acoustic sensors has been fused using a linear support vector machine (SVM) to achieve this goal. Main reason for choosing SVM is its popularity in classifying unlabeled features [85]. A description of the SVM approach and other key features that has led to use this method is followed.

Supervised SVM consists of a training phase and a testing phase. In the training phase, the model will be formed by mapping the ground truth data and optimizing a separating boundary between the labels. Here, assume the SVM is trying to label features in two classes of A and B using two features x and y. Consequently, in the training phase SVM calculates the maximum margin separating boundary. In the testing phase, SVM maps the feature vectors to see which side of the separating boundary they will be and
labels them accordingly. An illustration of the SVM training process is shown in Figure 22.

![Figure 22](image)

Figure 22. Training Process of Linear Supervised SVM with two predictors and two labels.

After the training phase, the model can be used for testing to evaluate its performance. The separating boundary extracted above will be used to label the new features after they are plotted.

Description above is for the simplest scenario where only two features are classifying the data into only two labels. In the cases where there are more than two features, the data will be mapped to a higher dimension. For example, if there were three features then the plot in Figure 22 would be three-dimensional, i.e. one dimension per feature. For the cases where SVM is aiming to label features in more than two categories, One-Against-One-approach will be used.

In One-against-one approach, multiple SVMs are fitted between each two labels. Consequently, each SVM will “vote” for a label. At the end, these votes are counted and the label with the most votes is chosen. An illustration of this method is shown in Figure 23 for three labels of A, B, and C and two features of x and y. As shown in the flowchart, each feature vector will be used for a separate SVM for classifying two labels. Afterwards, the label which was chosen more than other labels in the SVMs will be accepted as the final label [85- 87].
Figure 23. Illustration of one-against-one approach for three labels of A, B, and C and two features of x and y.

Experiment

Acoustic sensors’ parameters has been fused using a supervised linear Support Vector Machine (SVM) to predict features present on the road. Seven parameters from the sensors above were calculated for twenty-meter segments of surveyed roads. They were entered into data fusion model with labeled features of “cracks”, “manholes” and “no Features” known to be present in these segments from images of the camera on VOTERS vehicle. Forty-five segments were chosen, and thirty were used to train the SVM model. Figure 24 shows a graphical representation of this fusion model.

The SVM model was tested on fifteen twenty-meter sections with known features of Cracks, Manholes and No features. Figure 25 shows the prediction results for these segments.
Figure 24. Proposed fusion model for differentiating road features.

The fusion model performs well when there are cracks or no features present on the road. In case of manholes, however, the SVM model mislabeled them in three out of the five test sections. Possible reason could be that the tire did not hit those manholes properly, in the train or test sections. Future research to improve the fusion model’s performance in identifying manholes on the road is needed.

Figure 25. SVM Results for Feature Identification.
**Decision Trees and Priority Matrices for Suggesting Maintenance Activities**

After pavement condition evaluation, next step of a PMS process is identifying the maintenance activities required to restore pavements’ serviceability. In general terms, PCI-based systems can easily identify network-level maintenance categories. Typically, pavements below a PCI of 40 require reconstruction due to the significant structural damage. Pavements with a PCI of 40 to 70 can benefit from rehabilitation activities like overlay. If the pavement has a PCI score of above 70 and has not been damaged from load-related distress, it can be recover by performing preventive maintenance like crack sealing. [13, 15]. These numbers might be slightly different for arterial, residential and rural roads. A PCI value does not provide information on structural integrity of a pavement, which is where subsurface measurements will come in handy. Each of the stated types of repairs (preventive, rehabilitation and reconstruction) includes numbers of alternatives. PCI alone cannot zero-in on the optimal treatment among the available options. Decision trees are heavily employed by pavement management systems to make these project-level suggestions.

A decision tree is a type of classifier which classifies the dataset using a classification structure of the given problem and is composed of root nodes (zero incoming edges, zero or more outgoing edges), Internal Nodes (one incoming edge and should have two or more outgoing edges) and Leaf Nodes (one incoming edge and zero outgoing edges) [52, 53]. Each internal node denotes a test on an attribute, each branch represents an outcome of the test and leaf nodes represent the classes or class distributions. An appropriate attribute selection measure, partition strategy and stopping criteria are required for assembling decision trees.

A widely-used PMS with one of the most comprehensive decision trees is PAVER. PAVER uses PCI methodology and inventories seven shortlisted pavement distresses (twenty original ones) for each inspected road [54, 55] PAVER uses three severity levels for each distress and suggests a number of repairs that are specific to that severity level in each repair category of preventive, rehabilitation and reconstruction. VOTERS parameters can be used to provide these severity levels so that a decision tree specific to the VOTERS
system employing PAVER’s M&R suggestions. These severity levels for VOTERS parameters are provided in table 8.

Table 8. Severity levels for VOTERS parameters.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Low Severity</th>
<th>Medium Severity</th>
<th>High Severity</th>
</tr>
</thead>
<tbody>
<tr>
<td>MTD (mm)</td>
<td>&lt; 1.04</td>
<td>1.04 – 1.88</td>
<td>&gt; 1.88</td>
</tr>
<tr>
<td>IRI (mm/km)</td>
<td>&lt; 4</td>
<td>4 – 6</td>
<td>&gt; 6</td>
</tr>
<tr>
<td>Rutting Depth (mm)</td>
<td>&lt;10</td>
<td>10-25</td>
<td>&gt;25</td>
</tr>
<tr>
<td>Total Crack (mm/mm²)</td>
<td>&lt; 6</td>
<td>6 – 19</td>
<td>&gt; 19</td>
</tr>
</tbody>
</table>

Using the thresholds above, each parameter in the network-level category can be repaired by certain M&R activities. Once these probable repairs are determined they can be used to create appropriate attribute question, partition strategies and stopping criteria to find the dominant repair strategy. After thorough literature review of the latest maintenance techniques and their decision trees [56- 59], these alternative repairs for the three severity levels in table 8 have been defined. Consequently, different repairs will be suggested based on severity levels of each parameter out of which the mutual ones are chosen. Often two or three alternatives are available at the end of this process and one should be selected amongst them. In the reconstruction category (PC < 40), this selection is made by subsurface information and how solid the foundation of pavements are. In preventive and rehabilitation categories, traffic counts and date of the last repair will determine whether or not to choose the strongest repair option. Seasonality can also be a factor, e.g. depending on whether the repairs take place before or after a rainy season the M&R suggestions might be different.

Figure 26 summarizes PAVEMON’s M&R decision tree.
Figure 26. PAVEMON’s Repair Suggestion Decision Tree.

**Weight Matrices for Priority Assessment**

Maintenance requirements for restoring roads’ serviceable status often outpace available funding. Proper assessment of the priorities of these activates is required to optimize sending of maintenance budget. Current pavement management systems use three major strategies for prioritizing repairs.

- **Worst Repair Method**: Streets with the worst repair have the highest priority.
- **Least Cost Repair Method**: Streets with the lowest repair cost are given preference.
- **Benefit to Cost Ratio Method**: Streets with the highest Benefit to Cost (BCR) ratio are given preference. This ratio summarizes the overall value for money of a project, and takes into account the amount of monetary gain realized by
performing a project versus the amount it costs to execute the project. The majority of PMSs use this method [15, 54].

While these methods generally prioritize maintenance needs based on the pavement condition and repair cost, other decisive factors are in place that might not allow implementing the suggested strategy. For instance, an arterial road in a mediocre condition generally requires more maintenance consideration than a residential road in a very poor condition, or a bridge which is about to collapse due to subsurface deboning must be given the highest priority to impede catastrophic failures. Impacts of these parameters should be customizable to be geared more towards each agency’s unique requirements and goals.

PAVEMON uses a mathematical model to calculate these priorities. This model includes a priority matrix which weaves together VOTERS and third-party data in addition to a weight matrix with the weights (impacts) for each parameter. Final output gives the priorities of each repair indicating where the maintenance budget should be most directed towards.

Below is PAVEMON’s priority assessment model. All the parameters in this model has been normalized so that values measured on different scales have a notionally common scale.

\[
\begin{pmatrix}
N_{(100 - PCI)}_1 & N_{AADT}_1 & N_{AADT}_1 & N_{Jur}_1 & N_{Fou}_1 & N_{BCR}_1 \\
N_{(100 - PCI)}_2 & N_{AADT}_2 & N_{AADT}_2 & N_{Jur}_2 & N_{Fou}_2 & N_{BCR}_2 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
N_{(100 - PCI)}_n & N_{AADT}_n & N_{AADT}_n & N_{Jur}_n & N_{Fou}_n & N_{BCR}_n \\
\end{pmatrix} \cdot \begin{pmatrix}
W_{PCI} \\
W_{AADT} \\
W_{Age} \\
W_{Jur} \\
W_{Fou} \\
W_{BCR} \\
\end{pmatrix} = \begin{pmatrix}
Priority1 \\
Priority2 \\
\vdots \\
Priorityn \\
\end{pmatrix}
\]

Where

- \( N \) : Normalized parameters
- \( PCI_i \) : Pavement Condition Index of street \( i \)
- \( AADT_i \) : Average Annual Daily Traffic of street \( i \)
- \( Jur_i \) : Jurisdiction of street \( i \) (state, city, private, etc.)
\( Fou_i \) : Foundation reliability of street \( i \)

\( BCR_i \) : Benefit-to-Cost Ratio of repairing street \( i \)

\( W_X \) : Weight of parameter \( X \)

\( Priority_i \) : Priority of repairing street \( i \)

Normalization of parameters cited above are calculated from the equation below, which bring all values into the range (0, 1)

\[
N_X = \frac{(X - X_{min})}{(X_{max} - X_{min})}
\]

Where:

\( N_X \) : Normalized \( X \)

\( X_{min} \) : Minimum of \( X \)

\( X_{max} \) : Maximum of \( X \)

**Deterioration Model Considering Climate, Traffic, and Occurrences of Extreme Weather Events**

Predicting future budget needs and network condition needs an accurate deterioration model to quantify the effect of interactions between climate, vehicles and the road. Predicting this behavior is not easy. While deterioration models for rigid pavements have had a decent performance, because of the high visco-elastic characteristic of the asphalt, current deterioration models for flexible pavements have had limited success so far. Pavement infrastructure deterioration is an aggregated impact from traffic loading, environmental condition, and other contributors. The behavior of a pavement under these factors depends on the characteristics of its structure (materials and thickness of each pavement layer), the quality of its construction, and the subgrade (bearing capacity and
presence of water) [60]. Each factor causes certain distresses on the pavement. Understanding factors that lead to deterioration of roads help infrastructure managers to refine their construction and maintenance specifications. Following factors are known to be the main reasons for pavements’ degradation.

**Load:** Cracking and rutting caused by pavement bending under traffic loads are two of the most prominent forms of distresses. Tire pressure produced by vehicles in the radius of loaded area induces tensile stress on the pavement, lateral shear in the surface and vertical stress at the subgrade which gradually deteriorate the pavement [61].

**Material Properties:** Severity of distresses and the pace of their formation is heavily influenced by material properties of the pavement. Strength and bearing capacity, gradation, modules of elasticity and resilience of the materials used in construction determine pavements endurance under load and climate fluctuations. [62].

**Construction Quality:** Freitas et al. [63] shows that construction quality influences the two significant factors in initiation of top-down cracking, voids and aggregate gradations. Construction quality also determines the initial pavement condition which has an impact on the pace that pavement failures occur.

**Environmental Conditions:** Climate oscillations, precipitation and freeze/thaw cycles are the primary causes of some dominant distresses such as longitudinal and transversal cracks [64].

- **Temperature:** Temperature fluctuations are followed by tensile and compressive stress in pavement which initiates thermal cracking. Smith et al [65] shows a correlation between pavement deterioration and temperature where surge in temperature facilitates rutting and cracking in the pavement.
- **Precipitation:** Studies on pavement performance evaluations show that other than formation of longitudinal and alligator cracks, roughness of the road also worsens with a boost in precipitation.
- **Freeze/thaw cycles:** In cold regions, water penetrated into the pavement layers freezes in the winter. Thaw of these ice particles during spring causes deformation in pavement layers and triggers fatigue cracking [66].
Types of Deterioration Models

Depending on how aging of the pavement is simulated, road deterioration models can be categorized into deterministic and probabilistic models. Deterministic models are data driven mathematical functions typically trained with large amounts of datasets measured over a long period of time. Using these mathematical functions, these models predict future road conditions as a single value. Probabilistic models, on the other hand, provide a range of possible outcomes with the probability of their occurrences. These models are also referred to as Markov prediction models. Although considerable effort has been devoted to improve the quality of the probabilistic modeling of pavement deterioration, the applicability of their transition matrix is limited to only several widely spaced categories typically classified by traffic volume, pavement structure and climate regions [67]. Additionally, the fact that these models are used discrete in time led into adapting a deterministic approach by PAVEMON.

Deterioration due to Natural Causes

The Long-Term Pavement Performance (LTPP) program was established to collect pavement performance data and investigates pavement related details which are critical to pavement performance since the late 1980s. Over 2,500 test sections on highways throughout North America are monitored by LTPP. Following seven modules are measured: Inventory, Maintenance, Monitoring (Deflection, Distress, and Profile), Rehabilitation, Materials Testing, Traffic, and Climatic. Now that LTPP database contains more than two decades of data, valuable insights can be extracted from studying it.

Using provided data from LTPP, Jackson et. al. [66] preformed a multivariate regression analysis to predict pavement deterioration in terms of serviceability. They considered the following factors in the analysis:

- Pavement types (rigid, flexible).
- Climate (Precipitation, Cooling Index, Freezing Index, and thawing index).
- Stresses and strains calculated from layer material properties.
- Performance data (IRI).
- Soils and material properties.
Traffic data

Predicted performance measures were presented for each of the climatic scenarios and compared at a 95 percent confidence interval to determine statistically significant performance differences. Jackson et. al. [66] then derived an equation for both rigid and flexible pavements. As more than 85% of the roads in the United States are flexible pavements, the focus of the study here is flexible pavements, see equation below:

\[
\ln(\text{IRI} - 1) = \text{Age}(4.5\text{FI} - 1.78\text{CI} + 1.09\text{FTC} - 2.4\text{PRECIP} - 5.39 \log(\text{ESAL})/\text{SN})
\]

Where:

\[\Delta\text{IRI} : \text{Change in International Roughness Index}\]

\[\text{Age} : \text{Pavement Age}\]

\[\text{FI} : \text{Freezing Index} \text{ (Degree-days when air temperatures are below and above zero degrees Celsius)}\]

\[\text{CI} : \text{Cooling Index} \text{ (Temperature relation to the relative humidity and discomfort)}\]

\[\text{FTC} : \text{Freeze-thaw Cycles}\]

\[\text{PRECIP} : \text{Precipitation}\]

\[\text{ESAL} : \text{Equivalent Single Axle Load} \text{ (Conversion of traffic into single axle load)}\]

\[\text{SN} : \text{Structural Number}\]

Using large amounts of data for a long period of time in addition to the high correlation rendered at the end indicates the achievements of this model. However, this model still does not consider an important contributor in pavement deterioration: Effect of Extreme Conditions on road pavements. The devastation of New Orleans caused primarily by the breach of a levee during hurricane Katrina, the impact of hurricane Sandy on New York and New Jersey, a 16% immediate drop in road conditions of Denver in Colorado due to severe snow storms in 2006 are some examples that highlight the drastic effect severe events could have on pavements [69, 70]. The purpose here is to quantify the contribution
of two most prevailing events on pavement deterioration: Snow Storms and floods. These events exacerbate road conditions by causing shear failure and cracking, weakening the subgrade and widening the existing cracks, mainly due to the drastic increase in moisture content they cause in the pavement layers.

**Deterioration due to Extreme Weather Events**

Two consecutive IRI values of LTPP sections are measured one to four years apart and in irregular intervals. As no continuous data is available that entails IRI values of before and after a severe event, the effect of extreme events had to be quantified first.

By applying equation (1) and project one measured IRI to the point where the next IRI is measured, the two values should be reasonably close unless:

- Extreme events such as floods and snow storms occurred in that period. This causes the projected IRI values to be lower than the measure values.
- Maintenance and rehabilitation activities took place in that period. This causes the projected IRI values to be higher than the measured values.

To quantify the effect of extreme events on road deterioration, by forward-projected IRI values from one measured IRI to the point where the event has occurred. Consequently, the next measured IRI was backward-projected to the point where event has occurred; the difference between these two values is due to the extreme event that happened in that month if no other events/maintenance activities had taken place in that period. This is illustrated in Figure 27. To fit a model to these quantities, the required data had to be collected first.
Figure 27. Illustration of how the increase in IRI due to extreme events has been quantified using the immediate available IRI measurements.

**Data Collection**

In addition to the occurrence date of extreme events and the parameters that would define their magnitude and effect, all of the variables in LTPP’s natural deterioration model had to be collected. The first was acquired from LTPP database and the latter from National Oceanic and Atmospheric Administration (NOAA) database. Datasets were collected from January 1996 to December 2013 for the states of Florida, New Jersey, Ohio and Illinois. These states had the most comprehensive datasets available on LTPP and are more susceptible to frequent snow storms and floods.

- **LTPP Data Collection:** Most of the parameters were given in an annual format (e.g. FTC, FI, etc.) in the LTPP database. To isolate effect of an extreme event from natural deterioration, all datasets has been transformed to a monthly format. This
process involved interpolations and further calculations for some of the parameters, each had to be dealt with individually according to their meaning. The LTPP database lacked data in some places, some considerations and calculations had to be made, such as calculating cooling indices based on daily temperature, assigning missing SN values based on a study of pavement type and structural number of sections in that region and calculating missing ESAL values from available daily traffic and axle loads/numbers of vehicles.

- **NOAA Data Collection:** NOAA is a scientific agency focused on the conditions of the oceans and the atmosphere [71]. This database contained information on Snow Storms and Floods, their date of occurrence, duration and accumulated depth of water or snow.

  Using data from these four states, most of the time more than one event occurred between two consecutive measured IRIs. Only the data points were consider where the extra increase in IRI was due to one event in order to calculate the parameters of the model more realistically. Figure 28 shows the data collection and filtering process. Consequently, approximately 50 data points were identified for each individual event from around 4,000 data points.
Other than magnitude of the extreme event in terms of depth and duration, traffic (ESAL) and IRI values at time the event occurred were considered as predictors of deterioration. While each of these parameters provides worthy knowledge about this deterioration process, none alone can furnish sufficient information that will entail all needed to calculate the effect of Snow Storms and Floods. Our purpose here is to predict an increase in IRI from the right combination of all these parameters. Each of these events has been studied individually.
Table 9. Correlation of Predictors with increase in IRI due to the extreme event

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Correlation with %ΔIRI for Snow Storms</th>
<th>Correlation with %ΔIRI for Floods</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial IRI</td>
<td>0.721</td>
<td>0.317</td>
</tr>
<tr>
<td>ESAL</td>
<td>0.163</td>
<td>0.033</td>
</tr>
<tr>
<td>Event Depth</td>
<td>0.046</td>
<td>0.089</td>
</tr>
<tr>
<td>Event Duration</td>
<td>0.015</td>
<td>0.383</td>
</tr>
</tbody>
</table>

Results

Snow Storms: Stepwise regression was applied using all the four parameters shown in Table 9 in addition to four combinations of them. From the 42 available sections, 29 were entered into the fusion model along with all the potential predictors and response variables calculated for these streets to train the model. These sections were handpicked for training as they entail a diverse range of the predictors and response variable in order to fine-tune the fusion model’s boundaries. The fusion model used is shown in Figure 29.

Figure 29. Illustration of fusion model for quantifying the effect of Snow Storms on IRI.

In the final model, only four parameters were sufficient to predict IRI values, and the presence of any other parameter with the existence of these four was trivial, if not
harmful, to the correlation rendered by the model. The final equation derived through the fusion model is:

\[ \%\Delta IRI = 5.1 - 2.5 IRI_0 + 1.7 \text{ Depth} - 1.7 \text{ Duration} + 0.7 \text{ ESAL.Duration} \]

Where:

\%\Delta IRI: Percentage increase in IRI due to the snow storm

NIRI: Normalized IRI of the section before the snow storm

NDepth: Normalized Depth of the snow storm

NDuration: Normalized Duration of the snow storm

ESAL: Equivalent Single Axle Load (derived from traffic)

**Floods:** Similar to what was discussed above, eight parameters were entered into a stepwise regression model for 28 sections affected by a single flood, the remaining 7 sections were used for testing. The final equation derived through the fusion model is:

\[ \%\Delta IRI = 10.7 - 1.7 IRI_0 + 7.3 \text{ Depth} - 2.1 \text{ Duration} + 14.3 \text{ Depth.IRI} \]

Where:

\%\Delta IRI: Percentage increase in IRI due to the snow storm

NIRI: Normalized IRI of the section before the snow storm

NDepth: Normalized Depth of the snow storm

NDuration: Normalized Duration of the snow storm

ESAL: Equivalent Single Axle Load (derived from traffic)

IRI values of the remaining sections were predicted with the models developed above. Results were promising, rendering correlations of more than 90% for both events as shown in Figure 30.
Figure 30. Test Results for Floods (Left) and Snow Storms (Right).

As the main performance measure of PAVEMON is PCI, performance measures used in these equations have been converted to PCI using a correlation equation developed in the study by Park et al. [81]. The final Deterioration model is the following.

\[ \Delta PCI = \Delta PCI \_N + \Delta PCI \_S + \Delta PCI \_F \]

Where:

\( \Delta PCI \) : Change in Pavement Condition Index

\( \Delta PCI \_N \) : Change in Pavement Condition Index due to natural deterioration

\( \Delta PCI \_S \) : Change in Pavement Condition Index due to occurrences of snow storms

\( \Delta PCI \_F \) : Change in Pavement Condition Index due to occurrences of floods

Deteriorations cited above are quantified from equations below. Note that all the parameters are normalized before entering into the equations.

\[ \Delta PCI \_N = \text{Exp}(K \cdot \text{Age}(4.5 \cdot \text{NFI} + 1.78 \cdot CI + 1.09 \cdot FTC + 2.4 \cdot \text{PRECIP} + 5.39 \cdot \log(\text{ESAL})/\text{SN}) + 5.61)) \]

\[ \Delta PCI \_F = \text{Exp}(K (15.31 - 1.66 \cdot PCI_0 + 7.30 \cdot \text{Depth} - 2.10 \cdot \text{Duration} + 14.3 \cdot \text{Depth} \cdot \text{PCI}_0)) \]

\[ \Delta PCI \_S = \text{Exp}(K (9.70 - 2.5 \cdot PCI_0 + 1.7 \cdot \text{Depth} - 1.74 \cdot \text{Duration} + 0.706 \cdot \text{ESAL} \cdot \text{Duration})) \]
Where:

ΔPCI : Change in Pavement Condition Index
Age  : Pavement Age
FI   : Freezing Index (Degree-days when air temperatures are below and above zero)
CI   : Cooling Index (Temperature relation to the relative humidity and discomfort)
FTC  : Freeze-thaw Cycles
Precip  : Precipitation
ESAL : Equivalent Single Axle Load (Conversion of traffic into single axle load)
SN  : Structural Number
PCI_0 : Initial PCI
Depth  : Depth of the event
Duration  : Duration of the event

Using large amounts of data over a long period of time in addition to the high correlation rendered at the end are the achievements of this model (32, 35). This model is implemented as a part of PAVEMON’s future budget needs model.

**An Iterative Model for Predicting Future Budget Needs**

A PMS should be capable of determining the budget requirements for meeting specified management objectives. Typical management objectives include maintaining pavements above a specified condition and eliminating major M&R requirements over a specified number of years PAVEMON addresses both of these requirements by combining its maintenance decision tree and priority assessment model with the deterioration model discussed earlier.
Figure 31 shows the steps of identifying the required fixed minimum budget for reaching a target PCI after \( n \) number of years. An initial budget will be assumed at first. Then, maintenance and prioritization models will be applied to extract the PCI of the network after implementing M&R activities in the first year. Then, using the deterioration model discussed above, pavements degradations will be estimated to predict the PCI of the network in the second year. As not all the information gathered by the VOTERS system is available in the second year, a rough version of maintenance and prioritization models will be applied to extract the PCI of the network after treatments of first year. This rough version does not zero-in on the exact treatment and assumes the average cost of the conflicting repair options as the treatment cost. This process continues all the way to year \( n \). If the PCI of network in the final year is significantly less than the target PCI, the process needs to be repeated with a higher initial budget and vice versa. The initial value has been assumed to be \( 1/n \) of the total budget required to restore all the pavements in the first year.

Figure 31. Future cost prediction model.
**Outlook**

While PAVEMON considers a broad range of variables in its decision making algorithms, future research should focus on developing an algorithm for suggesting the type of material to use in the final maintenance alternatives. This would vary from region to region and street to street. Cost, endurance and availability of the material in addition to the performance expected from the road are the decisive factors in this decision tree. Another improvement that can be done is transforming PAVEMON’s single treatment suggestions to multiple treatment strategies to takes into account the historical maintenance activates on a pavement before suggesting a new one.
Developing a suitable system for the PMS methodology outlined in part 1 requires a thorough understanding of how VOTERS system collects data.

**An Overview of VOTERS DATA**

Each subsystem within the VOTERS system contains one or more sensors. Raw sensors data as well as the processed data are recorded as streams, i.e. a set of data in the same format and semantic meaning. One sensor can produce multiple streams, e.g. one can be the raw unprocessed stream such as video images of the road surface and another stream may be processed from the raw stream and show an image of the detected cracks. While both stem from the same sensor, they are considered separate streams.

A Global Positioning System (GPS) receiver collects the location and time information of the vehicle every second. All VOTERS streams are time synchronized with this sensor. Outputs of VOTERS surveys include the following:

**Positioning Data**

Contains spatial location of the sensor mounted vehicle and is collected every second. As the only stream with geo-spatial columns, positioning data enable geo-referencing the measurements of other sensors.

**Sensors’ Raw Data**

Emanates from two types of domains. One is the time-triggered sensors (dynamic tire pressure sensor, microphone, axle-accelerometer, laser height) which collect data at 50 kHz and are stored every 30 seconds in binary files. Another is the distance-triggered camera (video) which collects images every meter. This stream is stored as JPG files.
**Processed Data Streams**

Includes pavement-related parameters such as International Roughness Index (IRI), Mean Texture Depth (MTD), Road Profile, and Crack Density which are processed from individual sensors.

**Fused Data Streams**

Includes streams that are processed from multiple sensors such as the overall condition ratings or corrections made on the processed data through temporal correlation with the measurements of other sensors, all of which were discussed in previous sections.

VOTERS system produces these streams from multiple domains consisting of an array of heterogeneous sensor systems. Table 10 gives an impression about the domains and their recorded data amounts assuming a constant speed of 100 km/h.

<table>
<thead>
<tr>
<th>Domains</th>
<th>Max Sensor</th>
<th>Min Trigger Interval</th>
<th>Size/ point [byte]</th>
<th>Data rate [GB/h]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positioning</td>
<td>1</td>
<td>0.2 s</td>
<td>4</td>
<td>0.0003</td>
</tr>
<tr>
<td>Acoustic Microphones</td>
<td>4</td>
<td>25 us</td>
<td>4</td>
<td>2.1</td>
</tr>
<tr>
<td>Dynamic Tire Pressure</td>
<td>2</td>
<td>25 us</td>
<td>4</td>
<td>1.1</td>
</tr>
<tr>
<td>Video Systems</td>
<td>1</td>
<td>1 m</td>
<td>1</td>
<td>467.4</td>
</tr>
</tbody>
</table>

Temporal correlation across data streams (and consequently spatial correlation) is determined by time stamping each sample with microsecond accuracy in time [72]. Spatial correlation is achieved globally by the GPS receiver and local geometrical relationship of sensors to an absolute reference point on the vehicle.

As it can be observed, versatility of outputs and diversity of data granularities are two of the main challenges associated with implementing a suitable application for front-end users to leverage the data effectively.
Third-party Data

In addition to the VOTERS data, pavement-related information spanning from climate data such as temperature and precipitation to load data such as traffic counts which all influence pavement deterioration rate could be useful for decision making processes. These datasets are made available and fetched from sources such as National Oceanic and Atmospheric Administration (NOAA) and Long Term Pavement Performance (LTPP) databases [71, 74]. Most of these datasets can be consumed directly from the sources’ servers without any further processing. Additionally, manually collected road condition ratings by other cities and engineering firms are made available and placed into PAVEMON’s database for comparison purposes. These datasets are used to improve decision making and spatial analysis capabilities.

PAVEMON’s Architecture

Geographic Information System (GIS)

Zhang et al. 2012 [74] defines a Geographic Information System (GIS) as “a system of software and hardware modules for capturing, storing, visualizing, manipulating and analyzing all kinds of geographical data to facilitate pattern recognition, knowledge discovery and decision making”. Since geographical nature of the road networks are in parallel with spatial analysis capabilities of GIS, GIS has been long used for enhancing pavement management operations. Many states and cities have been moving towards integrating their historical data into GIS which is accompanied with features such as graphically displaying pavement conditions of roads of a network [75]. Displaying results of database queries, statistics and analyses on a map along with flexible database editing through dynamic color-coding of roadway sections in a graphical map interface are some of the advantages of such integration. Having a spatial meaning attached to all the VOTERS measurements GIS seems to be an ideal test-bed for this application and probably the only tool that can productively deal with it.

Because of the limited data collected by the current road inspection methods, road condition GIS maps typically consider an entire street or from intersection to intersection
as one line segment and append all the road distresses along that street to this line [74, 75]. Data collected with the VOTERS prototype can provide road features every meter and every second that the vehicle is traveling. Moreover, the images supplied by the VOTERS camera, captured every meter, provide a challenge on how to represent and geo-tag those as the source of these images is not stationary, unlike images used in most GIS maps. Thus a much more complicated application than the typical GIS systems is needed to handle these datasets.

**Flex API**

Providing an easy access to data and functionalities to the front-end users will increase the user-base of an application. With the current availability of internet everywhere, having a web-based interface is a decent way for enabling a global access. Any backend server technology can be integrated with a Flex project using Adobe’s Flash Builder to allow for building mapping apps on the Flex platform and incorporating a wide range of mapping and GIS tasks [76, 77].

**System Structure**

PAVEMON’s structure is shown in Figure 32. Main developments of PAVEMON are made in the GIS Server which interacts closely with the File Server. A description of this structure is followed.

**File Server**

After each survey is completed, raw data are uploaded to the file server, then processed to individual streams. These data are then being accessed by the GIS Server.

**GIS Server**

In the GIS Server, large volumes of data are given spatial references and evolve into meaningful knowledge. “Data intake scripts” feed available streams into a commercial grade work station hosting GIS software, an Oracle database, a web adopter and the Flex API. Python Scripts which use the ArcGIS library are responsible for geo-referencing streams and spatial analysis while the Oracle database stores the streams that the GIS software refines. Then “Fusion scripts” perform Data Fusion, e.g. to calculate PCI or do
noise-removal. To expose data on the GIS Server to an external audience, a request forwarding and security technology Web Adaptor was implemented [78]. This allows GIS server to integrate with an existing web server. After these layers are published to the web by the web adapter, the Flex API consumes them to grant access, enable queries and spatial analysis capabilities to front-end users. Prompted by the user, PAVEMON references each layer out of the database.

Consequently, for every second and every meter that the vehicle is moving, different pavement parameters are collected and processed which need to be georeferenced and placed on the map. These parameters include pavement related features of roughness, friction and road profile. As there no intentions to visualize or analyze all the data of all locations at the same time, data are stored in layers. Having a layer-based design allows for a more stable application as users will be querying only a fraction of the Terabytes of data managed by PAVEMON in an area and a parameter of interest.

**Georeferencing, Mapping, and Visualization**

PAVEMON’s GIS module includes two key components. First, an Oracle database which uses geo-references as the primary means of indexing information. This database is populated with the VOTERS and third-party data discussed earlier. Second, PAVEMON integrates spatial analysis functions that incorporate statistical models. With
its vast array of functions executable through a web-browser, PAVEMON allows users to perform computations on large data groups and view relationships that would otherwise not be obvious which differentiates PAVEMON from non-GIS mapping tools. The way in which PAVEMON stores, represents and analyzes data in addition to its GIS spatial tools are intended to mirror the way information are used for pavement management process. PAVEMON data structures include vector and raster data.

**Vector Data**

PAVEMON uses three basic geometric or foundational elements in vectors data: points, lines, and polygons. Points, which are basically single vertices, represent road features and objects such as manholes and potholes. Lines, which are non-closed sets of vertices, represent linear measurements over a period of time or distance that the VOTERS prototype has traveled, such as roughness of road segments or overall pavement conditions in terms of ASTM standards such as International Roughness Index (IRI) and Pavement Condition Index (PCI). Polygonal areas or regions, which are closed sets of bounding vertices, are used to represent climate data which are the same in each county.

Except video data, other VOTERS sensors are time-trigged, i.e. every second a measurement is available. To be more consistent, measurements need to represent same distances independent of the vehicle’s velocity at time of data collection. Thus two types of segmentations have been implemented: intersection to intersection segmentation and twenty-meter segmentation.

**Intersection to Intersection Segmentation:** Time-triggered datasets were segmented into polylines between two intersections. To aggregate a parameter measures between all intersections of a road network, positioning data of the start and end points are compared with the locations of all the intersections. Any two intersections that a measurement falls into are averaged and segmented as one polyline. Often times, some of the measurements falls between two intersections. For these areas, the one second segment is broken into two parts, one of which falls into one polyline and the other into the next one. Data that were collected when the vehicle was stationary (e.g. behind a traffic light) has been filtered out.
\((X_1, Y_1) \leq (X_i, Y_i) \text{ or } (X_j, Y_j) \leq (X_2, Y_2)\)

\(d_i = \text{MIN} \{V_1, \text{dis}(1,i), \text{dis}(2,i)\}\)

\(m_{\text{ave}} = \frac{(m_1.d_1 + m_2.d_2 + m_3.d_3 + \cdots)/(d_1 + d_2 + d_3 + \cdots)}{\sum m_i d_i / \sum d_i}\)

Where:

i and j : start and end of sensor data

1 and 2 : start and end of an intersection

V : velocity

m : measurement of the sensor

dis : distance between two points

Distance calculations cited above are calculated from the following:

\(a = \sin^2\left(\frac{\Delta Y}{2}\right) + \cos(Y_1).\cos(Y_2).\sin^2\left(\frac{\Delta X}{2}\right)\)

\(c = 2.\tan^2\left(\sqrt{\alpha}, \sqrt{(1-\alpha)}\right)\)

\(d = R.c\)

Where:

X and Y : longitude and latitude

R : earth’s Radius

**Twenty-meter Segmentation:** Time-triggered datasets were also segmented into 20 meter intervals. The uniform segmentation helps in making well-informed decisions. 20 meter was chosen as it showed the least noise and the most confidence while averaging the data
through correlating the results with the ground truth datasets acquired from local engineering firms.

**Raster Data**

NCHRP 2004 [75] defines Raster structures as “… dense arrays of values that represent features requiring large storage capacities and a lower nominal spatial resolution (byte). Processing raster data involves element-wise calculations.” As an example of raster data, an explanation of how video images are represented follows.

**Video image pop-ups:** VOTERS camera captures images of the pavement every meter. These have been represented by points hyperlinked to the image location on the file server. The paths to each of these images are stored in the database which the hyperlinks point to. As the granularity of the video images are generally higher than the positioning data (video time stamps are in nanoseconds), their location will be found by interpolating between the two nearest GPS points.

**Mosaic Layers:** Images taken by the camera behind the VOTERS van were stitched together and placed on the map where they were collected. This required finding the position and exact dimension of the images along with some distortion corrections. First, the images were projected from the image space to the coordinate space used by PAVEMON. This requires the dimensions and calibration of images, i.e. how many meters does each pixel equate into (Figure 34).  

\[
x' = AX + BY + C \\
y' = DX + EY + F
\]
Where:

X : number of columns in pixels (image space)

Y : number of rows in pixels (image space)

x' : length of the image in meters (coordinate space)

y' : width of the image in meters (coordinate space)

A : width of each pixel in meters

B : rotation term

C, F : coordinates (x and y respectively) of the upper-left pixel (could be any other one)

D : rotation term

E : height of each pixel in meters

Figure 34. Projecting images from image space to coordinate space.

The rotation terms defined above are determined based on vehicle’s direction at the time of collecting the image. This information is stored in the available GPS data. An example of how image rotations terms change with vehicle’s inclination is provided in Figure 35.
Figure 35. Example of how image should be rotated based on the vehicle’s inclination.

After images are stitched together, their pixel values are stretched over the same minimum and maximum values. By stretching them to the same RGB thresholds the images integrate seamlessly regardless of the sun settings or their exposure levels. An example is shown in Figure 36.

Figure 36. Mosaicked images before stretching (Left), and after stretching to the same RGB thresholds (Right).

Each image occupies around 2MB of space thus these images were published in the pyramid scheme to avoid caching lots of memory. Pyramids contain down-sampled layers of an original raster and expedite the display by retrieving a specified display resolution. Consequently, as one zooms in area required for displaying decreases while a higher-resolution copy of the data is displayed. Therefore, the performance is not affected
due to the trade-off between area and resolution. Figure 37 from [79] shows the pyramid scheme concept.

![Figure 37. Down-sampling pixel values in pyramid schemes.](image)

An example of the mosaicked layers discussed above is shown in Figure 38.

![Figure 38. Mosaicked layers of a street in Brockton, MA](image)

**Database Configuration**

PAVEMON utilizes an Oracle database which uses geo-references as the primary means of indexing information. This database is populated with the following VOTERS and third-party data:

- **Road Identifiers:** Includes street names, route numbers, width of the roads and lane numbers in addition to spatial coordinates of their start and end points.
- **Pavement Condition Evaluation Data:** Entails VOTERS pavement performance measures discussed earlier.
• **Traffic Data:** Contains traffic counts in terms of Average Annual Daily Traffic (AADT) and traffic loads in term of Equivalent Single Axle Loads (ESAL) which are obtained from the Department of Transportation (DOT) and Long Term Pavement Performance (LTPP) databases respectively [73].

• **Climate Data:** Precipitation, Freeze-thaw Cycles and other climate parameters which influence pavements’ deterioration rates are obtained from National Oceanic and Atmospheric Administration (NOAA) [71] server. These also include historical data of extreme weather events occurrences which will be used in PAVEMON’s deterioration model.

• **Cost of repairs:** Includes unit costs of the repairs suggested by PAVEMON and are obtained from the DOT websites and other local sources.

All of the information carries a spatial component to facilitate relating them. These datasets are utilized in the decision making models outlined in part 1.

**Tools and Functions**

PAVEMON functions include GIS spatial analysis capabilities such as measuring, charting and spatial queries in addition to database management capabilities of querying and statistical functions. Custom functionalities, such as incorporating Google Street-view into PAVEMON, have also been developed.

In Figure 39, the developed GIS visualization portal is shown with multiple data layers covering a given roadway. Multiple data visualizations are available such as crack density or the pavement image. In the query interface at the bottom, the user can fine tune the refined survey results such as the examination of single data point or a geographical area. In addition, a zoomed-in image embedded in the figure shows the detailed roadway performance validated by the images collected by the camera. Furthermore, Google Street-view has also been embedded into PAVEMON as a third party API. For each pin dropped on the map, Google Street View Panoramas will appear in a popup window as shown in top right of Figure 39.
From a front-end point of view, two aspects of PAVEMON are important to assess the value of this system: Data Layers and Functionalities.

**PAVEMON’s Data Layers**

An overview of the information made available by PAVEMON is followed.

**VOTERS Data**

**Positioning Data:** GPS sensor on top of the van gives location of sensor mounted vehicles every second. These are represented as points and polylines in several ways to show the path, speed, and heading of the vehicle at time of survey. An example is Figure 40 where speed (left) and heading (right) of the vehicle has been visualized.
**Raw Data:** Sensors’ raw streams, which are collected every 5 µs (at 200 kHz) and stored every 30 seconds in a binary files, can be downloaded through PAVEMON for further analysis and processing. These data are hyperlinked to their locations on the file server.

**Processed Data:** Processed from individual sensors represent pavement-related features of roughness, friction and distresses, parameters such as MTD, IRI are stored and represented for every second of data collection in addition to every 20 meters of the road. Figure 41 is an example of PCI layer which gives a rating every twenty-meter of the road, along with the other information such as information about pavement’s roughness (IRI) and texture (MTD). Grey, Red, Yellow, and Green represent Failed, Poor, Fair, and Good pavements respectively.

![Figure 41. Pavement condition parameters access and visualization.](image)

**Video Images:** Captured every one meter, images from the camera behind the van are accessed and visualized via PAVEMON through Pop-ups and Mosaicks, as shown in Figure 42. Currently, PAVEMON contains 1,300,000+ images.
Figure 42. Image Pop-ups for every one meter (Left), Mosaick layers of images (Right).

**Fused Data:** These streams are computed from multiple sensors’ raw and processed streams. Examples are Overall ratings in terms of pavement condition Index (PCI) and Road features. Example of road features in PAVEMON is shown in Figure 43.

![Figure 43. Road features visualized by PAVEMON, red symbols represent Cracks, Green symbols are patches and silver circles represent manholes.](image)

**Third-party Data**

**Manually Collected PCI:** Provided by local cities and engineering firms, these layers are used as a reference to fine-tune fusion algorithms. Figure 44 is a graphical representation of this layer on PAVEMON for the City of Brockton, MA.
Figure 44. Third-party PCI Data for Brockton, MA. Grey, Red, Yellow, and Green represent Failed, Poor, Fair, and Good pavements respectively.

**Reference IRI:** Provided by massDOT, this layer is used as a reference to fine-tune IRI calculated from DTPS and Accelerometer algorithm.

**Traffic Counts and load:** Provided from MassDOT and LTPP database, this layer helps in decision making algorithms as well as individual studies. An example of traffic counts layer on PAVEMON in terms of Average Annual Daily Traffic (AADT) for an area in Boston, MA, is shown in Figure 45.

Figure 45. Traffic Counts for an area in Boston, MA, where green, blue, yellow and red colors represent low to high AADT values respectively.
**Climate Data:** Provided from LTPP and NOAA databases, this layer helps in decision making algorithms as well as individual studies. An example is shown in top left of Figure 46, where blue, green, yellow, and red indicate average temperature of -13 °C to 40 °C respectively.

**Extreme Events Data:** Provided from Federal Emergency Management Agency (FEMA) [83] and NOAA databases, this layer helps in deterioration model outlined before in addition to individual studies. Figure 46 contains examples of this layer from.

![Figure 46](image)

Figure 46. US Weather Forecast from NOAA (Top Left), Temporal Hurricane Tracks from NOAA (Top Right) and Flood Zones from FEMA (Bottom).

Table 11 shows a full list of layers on PAVEMON. Note the diversity of data, their shapes, segmentations, and sources.
Table 11. Full list of PAVEMON’s layers.

<table>
<thead>
<tr>
<th>Layer</th>
<th>Feature</th>
<th>Shape</th>
<th>Segments</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>FUSED</td>
<td>Pavement Condition Index (PCI)</td>
<td>Polyline</td>
<td>As small as 20meters</td>
<td>Microphone DTPS</td>
</tr>
<tr>
<td></td>
<td>Road Features (Cracks, Patches, Potholes, Manholes)</td>
<td>Symbol</td>
<td>Where feature exists</td>
<td>Microphone DTPS mm-wave</td>
</tr>
<tr>
<td>PROCESSED</td>
<td>Mean Texture Depth (MTD)</td>
<td>Polyline</td>
<td>As small as 20meters</td>
<td>Microphone</td>
</tr>
<tr>
<td></td>
<td>Roughness Index (IRI)</td>
<td>Polyline</td>
<td>As small as 20meters</td>
<td>DTPS</td>
</tr>
<tr>
<td></td>
<td>Crack Density</td>
<td>Piechart</td>
<td>Where feature exists</td>
<td>Camera</td>
</tr>
<tr>
<td></td>
<td>Road Profile</td>
<td>Polyline</td>
<td>User selectable</td>
<td>DTPS</td>
</tr>
<tr>
<td>VIDEO</td>
<td>Mosaic Layers</td>
<td>Mosaic</td>
<td>Where feature exists</td>
<td>Camera</td>
</tr>
<tr>
<td></td>
<td>Image Pop-ups</td>
<td>Point</td>
<td>Where feature exists</td>
<td>Camera</td>
</tr>
<tr>
<td></td>
<td>Cracks Statistics</td>
<td>Point</td>
<td>Where feature exists</td>
<td>Camera</td>
</tr>
<tr>
<td>POSITIONING</td>
<td>Vehicle Location</td>
<td>Point</td>
<td>Where feature exists</td>
<td>GPS</td>
</tr>
<tr>
<td></td>
<td>Vehicle Path</td>
<td>Line</td>
<td>Where feature exists</td>
<td>GPS</td>
</tr>
<tr>
<td></td>
<td>Vehicle Speed</td>
<td>Point</td>
<td>Where feature exists</td>
<td>GPS</td>
</tr>
<tr>
<td>RAW</td>
<td>Raw Binary Streams</td>
<td>Point</td>
<td>Where feature exists</td>
<td>All Sensors</td>
</tr>
<tr>
<td>NON-VOTERS</td>
<td>Manually Collected PCI</td>
<td>Polyline</td>
<td>Every street</td>
<td>CDM Smith FS&amp;T</td>
</tr>
<tr>
<td></td>
<td>IRI of MA Highways</td>
<td>Polyline</td>
<td>Every street</td>
<td>MassDOT</td>
</tr>
<tr>
<td></td>
<td>Traffic (AADT and Load (ESAL))</td>
<td>Polyline</td>
<td>Every street</td>
<td>LTPP Database</td>
</tr>
<tr>
<td></td>
<td>Climate and Extreme Events</td>
<td>Polygon</td>
<td>Geographic Regions</td>
<td>NOAA</td>
</tr>
</tbody>
</table>

Examples of PAVEMON’s Functions

Prominent functionalities of PAVEMON for the front-end users are discussed below.

Data Access and Query

Users have the ability of exporting features. They can draw a box on an area of interest and see the detailed features such as PCI, Latitude, Longitude and other parameters in a table. Users can then export those features as a text or CSV file for further inspection. PAVEMON also has the capability of letting users scroll quickly through image thumbnails in the attribute table of layers and look for a certain feature (pothole, cracks, patching, etc.) along with other parameters in the table (Figure 47).
Google Street-view and Birds Eye view

Google Street-view images have been integrated into PAVEMON as a third party API to further verify sensors’ data and help in algorithm development. With this capability, PAVEMON provides users with the ability of dropping Google’s “Pegman” on a map with sensors’ data and walk on roads while observing VOTERS data. Moreover, Birds Eye view can be seen along with these images, as shown in Figure 48.
Data Charting and Summary

Users can summarize any of the parameters in a selected area and derive a parameter vs. mile chart. Consequently, users can observe the changes of pavement parameters along a line and quickly identify the weak spots of the network. In Figure 49 crack densities has been summarized for a custom area on the map.

![Figure 49. Crack Densities plot for a custom selected interval.](image)

Help Documents and References

With the links and documents provided and accessed through PAVEMON, users can reference and understand the parameters of roads used in the map. Additionally, references on the sensors and their parameters are provided as shown in Figure 50.
Reports and Statistics

One of the prominent information roadway agencies and infrastructure managers want to know about their road network is the distribution of roads in each PCI categories (Failed, Serious, Fair, Satisfactory, etc.). This is essential for top level budget planning and funding proposals. As these processes has to be one the fly, not hard coded, an endeavor for development of this functionality was undertaken. Currently, two kinds of statistics can be extracted easily through PAVEMON for the whole network or a user-defined custom area:

1. Percentage and mileage of roads that fall above and below critical levels.
2. Distribution of road conditions in the seven PCI categories in mileage and percentage.

These can be presented in Pie-Chart or Bar-Chart formats, as shown in Figure 51.
High-resolution Temporal Analysis

A key feature of the sensor system monitoring framework is the capability of collecting pavement-related data frequently. Having this valuable time-series dataset, the opportunity to study pavement degradation based on real detailed data becomes feasible.

A tool was developed to closely study how the cracks have spread out and how road conditions in term of PCI, IRI or MTD have deteriorated. With this tool, one can navigate through mosaicked images collected at different times and see how a certain crack has been progressed or how a pothole has been formed, as shown in Figure 52. This tool can be executed on any other layers such as PCI, MTD or IRI.

Figure 51. Distribution of road miles & percentage in seven PCI categories (Left), in two (Right).

Figure 52. Examining how cracks of a certain location have spread out in span of a year.
PaveMan Toolbox

Paveman toolbox has been developed to provide effective tools and methods that can assist decision makers in formulating optimum strategies for providing and maintaining a serviceable pavement network over a given time period (the planning horizon). PaveMan leverages VOTERS and third-party data for an organized data driven approach for agencies to conduct pavement management activities. This decision-support module uses the algorithms described in Part I of this synthesis.

Projecting PCI

Prediction modeling can show streets deteriorating over time. By using the data driven deterioration model that takes into account occurrence of extreme weather events described before, one can visualize road conditions looking back and forth in time with the time-slider embedded in PAVEMON. Figure 53 shows an example where the user has been using this tool to visualize the road network back and forth in time.

Maintenance Strategies

Maintenance Strategies tool consists of the following components:

Repair Suggestions: Effective treatments ranging from slurry seal to an overlay to a full removal and reconstruction are suggested using a decision tree discussed in Part I. With this tool, user can customize costs of each repair and choose to exclude certain repairs from the suggestion results and generate results by drawing an area on the map or simply choose
the whole network. Decision tree developed based on VOTERS diverse datasets and used for streets generating repair suggestions is discussed in part I of the thesis. Figure 54 shows the user selecting an area of the network and the system automatically generating the repair suggestions accordingly.

![Figure 54. PaveMan’s Repair Suggestion’s Tool.](image)

**Repair Prioritization:** With this tool, user can enter a budget and streets that can be repaired will appear on the map. In the settings of this tool one can change the default priorities and impose certain streets to be included in the results. User can interactively see how difference scenarios will change the network’s performance. Figure 55 shows an example ran with entering a custom budget. Blue segments are the streets that the system assessed can be repaired with the budget entered.

![Figure 55. PaveMan’s Repair Prioritization Tool.](image)
**Long-term Planner:** With this tool, the user defines a target PCI to reach after a user-defined number of years and the system will generate the estimated budget required based on the current and predicted road conditions. The user has also the option of entering capital investment amounts (one time improvement amounts) for each year, as shown in Figure 56.

![Long Term Planner](image)

**Figure 56.** PaveMan’s Long-term Planner.

**Settings:** PaveMan is fully customizable. Users can change default priorities and change the repair decision trees to fit his/her preferences. Users can also interactively change the settings to see the effect of different decisions.

![Repair Settings](image)
![Priority Settings](image)

**Figure 57.** PaveMan’s Settings
Help Pages: PaveMan has comprehensive help files linked to each tool. These files describe user inputs and default settings in addition to an operation manual. An example of these help files is shown in Figure 58.

![Figure 58](image)

Summary of PAVEMON’s Implementation Challenges

The nature of the VOTERS project and Terabytes of data pouring in from various modalities after each survey day pose many different data handling and management challenges. Some of these challenges include:

- Proper Geo-referencing of all data with different modalities
- Handling huge amounts of data (TB) in each survey
- Road features in increments as small as one meter
- Dealing with time-triggered and distance-triggered datasets
- Geo-referencing Images from the camera which are captured every one meter from a non-stationary source
Large amounts of data with diverse modalities: The VOTERS prototype can produce TB of data in a day. Other than size, different modalities of VOTERS datasets require unique approaches for each.

Considering that multiple such systems might collect data simultaneously it is crucial to automate the data flow through the system from the data acquisition to where it is automatically processed and placed into the database. By using python scripts and data models, the process inside GIS server is automated and can place data on PAVEMON within a day of the survey.

Performance issues because of large amounts of data: Representing large amounts of data would slow down the application as it will cache a lot of memory. Thus data are stored in layers to speed all the processes and analyses to prevent this issue from occurring. Another benefit of having layers is that the user doesn’t want to query or visualize all the data at once, it would be tough to make decisions and see patterns with all data visualized and represented at once. The user wants to do it per parameter and in a certain area.

Mosaick Layers: Finding Position and High memory cache of images were two of the main challenges of dealing with the hundreds of thousands of images. Mosaicking these images required finding the position and exact dimension of the images along with some distortion corrections. Furthermore, each image occupies around 2 MB of space, thus mosaicking them would occupy lots of memory. By publishing these images in the pyramid scheme this problem was overcome as described earlier.

Noise in time-triggered data: Acoustic data and fused streams are time triggered, i.e. every one second there are 50,000 data points (sampling at 50 kHz). Depending on speed of the vehicle, this one second segments differs in distance. A filter of 10, 15 and 20 meter segments has been defined for IRI, MTD and PCI values using Python scripts. It was concluded that when values were average for 20 meters, noise was a minimum and results were reliable. Thus processed and fused streams (IRI, MTD, PCI, etc.) were represented in 20 meter segments as well as from intersection to intersection.

Manhole and Speed effects: Often times that the vehicle steps on a manhole, it is misidentified as a road defect. Furthermore, high accelerations will cause uncertainty and noise
in data. These two problems were dealt with by identifying and removing the areas affected by these effects and reconstructing the shape by spatial averaging between two ends of the removed sections. This methodology was described in Part I: PAVEMON’s Design.
SUMMARY AND OUTLOOK

This thesis has introduced PAVEMON, an online GIS tool for accommodating, representing, and leveraging pavement information layers, most of which are emanating from a multi-modal mobile sensor system. The algorithms developed to assess pavement condition through temporal data fusion were presented, and they were followed by the decision making algorithms developed to assist the roadway agencies for making the most cost-effective maintenance schedules. Examples of PAVEMON’s interfaces and functionalities were given at the end.

At the time this thesis was written, PAVEMON contained around 1,300,000 images and 3 TB of sensors and third-party data. Users operate on these datasets and do different types of analyses through a web browser and without needing to install any software. PAVEMON has custom toolboxes and interfaces which have been created using a combination of Python and Action scripts. A decent example of these tools is PAVEMON’s Pavement Management Module which allows decision makers to do budget planning and maintenance scheduling from countless datasets managed by this application.
PAVEMON is also capable of displaying VOTERS vehicle in real-time as it is collecting the data. The many components of PAVEMON, which work in sync to provide the most up-to-date information to the decision makers, are shown in Figure 57.

**Outlook**

PAVEMON provides an excellent opportunity for uniform maintenance scheduling across organizations. Integrating all the city’s assets such as signs, pipeline and list of materials that endure the most in each region allows for synchronized maintenance activities across organizations. As an example, if some utilities are to be replaced where roads are about to be reconstructed, reconstruction should be delayed for after the utility replacement has been done. This would be possible if both entities operate on a common platform. Other than engineers and infrastructure managers, economic developers can benefit from such a platform. As an example, retail businesses study traffic counts and road conditions before choosing a new business location. Having PAVEMON as a common platform for all these information will have enormous applications. Figure 60 illustrates this vision.

![Diagram](image.png)

Figure 60. PAVEMON to weave together VOTERS and third-party data to leverage a city/state’s data investment.

Imagining PAVEMON to evolve as a global road condition app someday is intriguing. A web-map that allows municipalities as well as public to see their current
infrastructure conditions while vehicles carrying customized sensors are roaming around the city. Cities would then no longer need to impose road closures for inspections by having such detailed information at their fingertips. Considering the enormous benefits of having a sound road infrastructure, this portray might not be that costly to paint.
REFERENCES


[28] Shahini Shamsabadi S., Birken R., Wang M., Pavement Condition Monitoring by Fusing Data from a Mobile Multi-sensor Inspection System, 6th International Association for Structural Control and Monitoring (IACSM), July 2014


[64] Li W., Mills L., McNeil S., The Implications of Climate Change on Pavement Performance and Design, University of Delaware University, Transportation Center (UD-UTC), 2011.


[67] Li N., Xie W., Haas R., Reliability-Based Processing of Markov Chains for Modeling Pavement Network

[68] Deterioration, Transporation Research Record 1524.


APPENDIX: PAVEMON’S ROUTING SYSTEM

VOTERS vehicle is aiming to cover cities similar to how a truck would plow snow: full streets and all lanes must be covered. This is a challenging logistics solution, especially for large scale surveys. Therefore, great effort was made to develop a routing system in PAVEMON to navigate the VOTERS vehicle throughout the city, to collect data on the full length of roads and all lanes efficiently. PAVEMON’s routing system has two main objectives: 1) covering full length of roads, 2) covering all lanes.

1) Covering Full Length of Roads

PAVEMON automatically defines hundreds of waypoints every 100 meters on the streets of interest. Consequently, when least-cost network path are identified, the consecutive waypoints navigates the driver to hit each one resulting in a full length coverage of the road.

2) Covering All Lanes of Roads

Using TomTom’s street-map data, number of lanes for each road is known. Consequently, the following changes are made to the waypoints defined in the previous part based on the number of lanes each road has. The strategy is different for two-way streets, one-way streets, and dead-ends.

Two Way streets

Based on the number of lanes, following solutions were developed:

Streets with two lanes: For these streets, half the waypoints defined in the A are pushed to the right side of the road, and the other half to the left. Consequently, the driver needs to hit the waypoint going at each direction once to “drop off passengers at the right side of the road.”, as its optimization method is based on the assumption that these waypoints are drop-off locations for a school-bus.
**Streets with three-lanes:** For these streets, same as the streets with two lanes, half the waypoints defined in the A are pushed to the right side of the road, and the other half to the left and the driver is forced to hit the waypoints on the direction which has more than a single lane, twice. Consequently, the waypoints are named “right lane”, “left lane” so the driver knows which lane to drive when importing the routes on GPS navigation system.

**Streets with four lanes:** For these streets, same as the streets with two lanes, half the waypoints defined in the A are pushed to the right side of the road, and half to the left. If each direction has two lanes, the driver is forced to hit these waypoints twice. Consequently, the waypoints are named “right lane”, “left lane” so the driver knows which lane to drive each time when importing the routes on GPS navigation system.

If one direction has three lanes and one has a single lane, the driver is forced to hit the on the former direction three-times. Consequently, the waypoints are named “right lane”, “left lane”, “middle lane” so the driver knows which lane to drive each time when importing the routes on GPS navigation system.

As this routing developments was done for the city roads, highways with more than four lanes were not considered. However, PAVEMON’s routing system can be easily extended to consider highways with more than four lanes.

**One-way streets**

Based on the number of lanes, following solutions were developed:

**Streets with a single lane:** No changes are made to the location of waypoints and the driver only needs to hit them once.

**Streets with two lanes:** For these streets, the driver is forced to hit the waypoints twice. Consequently, the waypoints are named “right lane”, “left lane” so the driver knows which lane to drive when importing the routes on GPS navigation system.

**Dead-ends**

For each dead-end, one waypoint at the end is automatically defined to force the driver to survey these roads. Consequently, the dead-ends are named accordingly so the driver
knows in advance he needs to turn around at the end, as some of the GPS navigation systems get confused how to navigate to the next waypoint at dead-ends.

Figure 61 shows the routes defined for a network in Cambridge, MA, by PAVEMON. Blue and Purple circles are the waypoints automatically placed on the roads of interest with the framework outlined above.

![Figure 61. PAVEMON’s customized routing system that automatically defines hundreds of waypoints on the streets of interest and navigates the driver to the appropriate lane on each drive.](attachment:image.png)