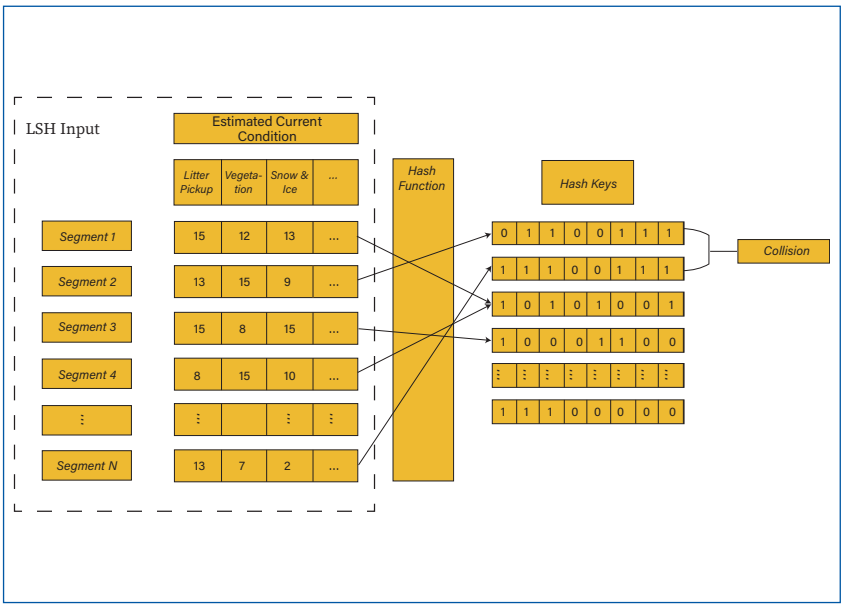
MOUNTAIN-PLAINS CONSORTIUM

MPC 19-392 | X.C. Liu and Z. Chen

Hotspot and Sampling Analysis for Effective Maintenance Management and Performance Monitoring



MOUNTAIN-PLAINS CONSORTIUM REGION 8

A University Transportation Center sponsored by the U.S. Department of Transportation serving the Mountain-Plains Region. Consortium members:

Colorado State University
North Dakota State University
South Dakota State University
University of Colorado Denver
University of Denver
University of Utah
Utah State University
University of Wyoming

Hotspot and Sampling Analysis for Effective Maintenance Management and Performance Monitoring

Xiaoyue Cathy Liu, Ph.D., P.E.
Assistant Professor
Department of Civil and Environmental Engineering
University of Utah
Salt Lake City, Utah, 84112
Phone: (801) 587-8858
Email: cathy.liu@utah.edu

Zhuo Chen, Ph.D.
Postdoctoral Researcher
Department of Civil and Environmental Engineering
University of Utah
Salt Lake City, Utah, 84112
Email: zhuo.chen@utah.edu

July 2019

Acknowledgements

The authors acknowledge the Mountain-Plains Consortium (MPC) and the Utah Department of Transportation (UDOT) for funding this research, and the following individuals from UDOT on the Technical Advisory Committee for helping to guide the research:

- Rukhsana Lindsey
- Lloyd Neeley
- Tim Ularich
- Jason Richins
- Tammy Misarasi
- Todd Richins

Disclaimer

ABSTRACT

A high-dimensional clustering-based sampling method for roadway asset condition inspection is proposed in this study. The method complements existing literature by selecting sample roadway segments that contain multiple types of assets (e.g., signage, shoulder work, pavement marking, etc.) for the accurate estimation of their respective levels of maintenance (LOMs). This is consistent with the standard maintenance procedure, as inspection activities are often conducted on roadway segment basis. The proposed method consists of three components: current condition estimation, similarity matrix construction, and stratification. Current condition estimation predicts assets' "current condition" by considering historical inspection records. Similarity matrix construction represents the core piece of the sampling framework, which employs a locality-sensitive hashing algorithm to define the similarity between segments. The stratification process is implemented with spectral clustering, which assigns segments into clusters based on the similarity matrix. The proposed method outperforms simple random sampling, which is widely used by state agencies, especially under the circumstances where LOM varies greatly across assets. The main highlight of the proposed method is the ability to select sample segments with multiple types of assets that are representative of their respective LOMs of the full inventory, which directly translates into an efficient maintenance activity management. The method is implemented using asset inspection records in the state of Utah from September 2014 to March 2016. It represents a potentially useful tool for agencies to effectively conduct asset inspection, and it can be easily adopted for choosing samples containing multiple features.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

EXECUTIVE SUMMARY

Asset field inspection is critical to effective performance monitoring and asset maintenance. Given the constraints on budget and time, the inspection activities thus require significant attention for planning and monitoring. Particularly, a sampling technique is needed for each asset item to determine the portion of population to be inspected, frequency at which the inspection should be conducted, and method to be used to collect the information. In asset management, DOTs usually use highway segment as the sampling unit, where more than one type of asset exists for inspection. Thus, it is time-consuming and operationally inefficient for field personnel if samples of different assets are widely distributed across segments. An ideal sampling method is to select a group of highway segments for inspection, in which sampled assets are representative to reflect their respective level of maintenance (LOM), a performance measure of maintenance activities, within the entire network. To this end, the project develops a high-dimensional clustering-based stratified sampling (HDCSS) method for roadway asset inspection. The sampling segments selected by this method can accurately represent the overall asset conditions of the full inventory. The HDCSS method generally outperforms the simple random sampling (SRS) method, which is widely used by DOTs. The method consists of three components: current condition estimation, similarity matrix construction, and stratification. The current condition estimation aims to provide predicted asset condition for clustering based on historical inspection records. Segments with multiple asset types are treated as high-dimensional vectors. By applying a locality-sensitive hashing (LSH) algorithm and spectral clustering, the similarity between segments is measured and the segments are assigned to clusters (strata). Using inspection records from the State of Utah, it is verified that HDCSS outperforms SRS for most asset types, especially under the circumstances where LOM varies greatly within assets. For the assets with little LOM variations across segments, both the HDCSS method and SRS yield low errors. Our proposed method requires a relatively smaller sample size, leading to a potential decrease in inspection costs, especially for large-scale networks.

By using the HDCSS method, DOTs can save resources and time for asset inspection, because inspection is carried out on the segment basis and the similarity identification introduced through an LSH algorithm. The method can be further applied to any high-dimensional sampling process, e.g., in selecting corridor segments, intersections, or traffic assets where multiple types of features, such as traffic condition, geometric design, or assets, need to be considered.

1. INTRODUCTION

1.1 Problem Statement

Asset management, often referred to as the decision-making process to allocate resources for roadway asset preservation (Durango-Cohen, 2007; Sousa et al., 2017), includes three major components: inspection, maintenance, and rehabilitation. Asset management agencies evaluate the current conditions of roadway assets, such as road shoulder, signage, and pavement marking, via inspection. Decisions regarding which prescribed maintenance and rehabilitation activities are then conducted, and how the transportation investments are prioritized can then be made. Inspection is thus critical, as it directly ties to the planning of maintenance and rehabilitation activities. Any inaccuracy in inspection will impair the reliability of maintenance and rehabilitation decisions. Yet, collecting asset condition data is very demanding in terms of labor and time. Oftentimes, agencies inspect only a portion of the assets, rather than the entire asset inventory to estimate the overall condition. As Mishalani and Gong (2009a) mentioned, there are four factors associated with the accuracy of inspection results: inspection frequency, inspection technologies and data processing methods, sample size, and correlation between observations. Three of the aforementioned factors (inspection frequency, sample size, and correlation between observations) are pertinent to the selection of sampling method. Improperly selected sampling method can be a major source of error, but when chosen appropriately it can be a useful tool for accurate condition estimation. Compared with other options for improving inspection accuracy, such as adopting advanced inspection technologies, the use of proper sampling methods improves inspection quality with minimal investment. Besides the accuracy of inspection results and initial investment costs, another concern about asset inspection is the recurrent costs. One common sampling method used by departments of transportation (DOTs) is simple random sampling (SRS), which selects samples based on a random draw. The method is unbiased and able to generate samples for any asset types simultaneously. However, it always requires a large sampling rate to justify the representativeness of samples. Another widely used sampling method is stratified random sampling, which divides a population into strata and selects samples from each stratum. It can use relatively smaller sample rates, yet the sample selection is only confined to a single type of asset at a time. In asset management, DOTs usually use highway segment as the sampling unit, where more than one type of asset exists for inspection. It is thus time-consuming and operationally inefficient for field personnel if samples of different assets are widely distributed across segments. An ideal sampling method is to select a group of highway segments for inspection, in which sampled assets are representative to reflect their respective level of maintenance (LOM), a performance measure of maintenance activities, within the entire network. Such sampling method, enabling once-for-all inspection instead of once-for-each-asset-type, has the potential to significantly reduce inspection costs. Different from quality control sampling or acceptance sampling where the previous condition of individual sample is unknown, asset management agencies have full access to historical records of highway asset conditions on the sampled segments, i.e., location, maintenance log, inspection results, and even latent risk estimation. It facilitates the design of an information-based sampling method that can select the most representative samples on the basis of historical records. Moreover, when historical information includes updated inspection results and maintenance records, effective sampling methods can always dynamically adjust the sample selection.

1.2 Objectives

The primary objective of this research is to develop a sampling method utilizing machine learning techniques to suggest the location and frequency of sampling roadway asset. The method should strive to choose proper segments where the conditions of sampled assets can represent the maintenance performance of the full inventory within the network.
To this end, we present a high-dimensional clustering-based stratified sampling (HDCSS) method for roadway asset inspection. The method allows transportation agencies (e.g., DOTs) to adjust parameters, such as sample size, inspection frequency, and assets of interest. The HDCSS method integrates asset deterioration prediction, high-dimensional clustering, and locality-sensitive hashing (LSH). The sampling method can also incorporate various features of the asset network, such as asset condition, geographic information, traffic condition, and geometric design, as the information upon which samples can be selected. The method is adaptable to any asset changes, as the sampling process is constantly updated with previous inspection results and maintenance records.

1.3 Outline of Report

The rest of the report is structured as follows. Section 2 summarizes the literature on asset sampling and high-dimensional clustering. The proposed HDCSS method and its mechanism are explained in Section 3. Section 4 demonstrates the application of our sampling method with asset data collected by the Utah DOT (UDOT) and artificial data generated via simulation, as well as results and analysis. Implications and conclusions are presented in Section 5.

2. LITERATURE REVIEWS

2.1 Review on Roadway Asset Management

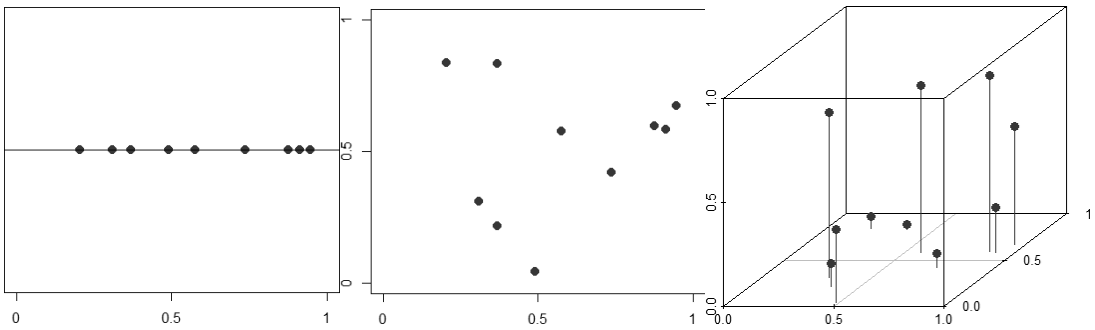
In the stream of research on roadway asset management, inspection is often jointly studied with maintenance (Guillaumot et al., 2002; Liu and Chen, 2017; Memarzadeh and Pozzi, 2016; Mishalani and Gong, 2009a, 2009b; Smilowitz and Madanat, 2000). One of the most widely used maintenance optimization algorithms is the Latent Markov Decision Process (LMDP). In LMDP, asset conditions are represented with a set of discrete states, and the deterioration process is encoded as Markovian transition matrix with probabilities. It considers uncertainty introduced by asset performance prediction and measurement (Ben-Akiva et al., 1993). Output of the method is the optimal inspection and maintenance policies. LMDP aims to either minimize total managing costs or maximize asset performance over finite/infinite planning horizon. Several efforts have been made to further refine the LMDP method since it was first proposed (Durango-Cohen and Madanat, 2008; Mishalani and Gong, 2009c; Smilowitz and Madanat, 2000). Smilowitz and Madanat (2000) extended the LMDP model to a network-level problem by including network-level constraints, e.g., minimum/maximum percentage of assets allowed in certain conditions and minimum/maximum yearly budget. Durango-Cohen and Madanat (2008) improved the LMDP method with adaptive control formulations. Instead of using a single Markov Decision Process (MDP) to represent asset performance and transition, their optimization model used a finite mixture of MDPs. Also, the asset deterioration process was constantly updated with a feed of new condition measurements. Mishalani and Gong (2009a; 2009c) incorporated uncertainties from sampling (i.e., sample size, spatial sampling) and the inspection process, and used them as decision variables in LMDP models. Maintenance activity on each segment (repair, inspection, or do nothing) was optimized based on consequences if such maintenance activities were conducted. In all these aforementioned approaches, the fraction of inspection (sample size) is determined as a result of optimization.

One classic sampling method in transportation maintenance management is stratified sampling, since it balances trade-offs between inspection costs and sampling accuracy (Adams and Winkelman, 2016; De la Garza et al., 2008; Medina et al., 2009). De la Garza et al. (2008) proposed a stratified sampling method for road maintenance evaluation. The stratification criteria include geographical location, weather variation, urban and rural setting, and traffic volume. Adams and Winkelman (2016) evaluated the effectiveness of a stratified random sampling method for transportation assets. They pointed out that by employing stratified sampling techniques rather than the SRS method, agencies can reduce sample size and improve precision. Medina et al. (2009) proposed a sampling protocol that stratifies the population with functional classification, annual average daily traffic (AADT) range, and asset category. In most of these previous studies, segments were stratified based on features of road segments (geometric design, functional classification, AADT) rather than the conditions of assets. In such cases, the stratification sampling method may under- or over-estimate the asset condition when the variance of asset LOM is high.

2.2 Review on High-Dimensional Clustering

In HDCSS, stratification is implemented via high-dimensional clustering. Since each highway segment often contains multiple types of assets, we consider a segment as a high-dimensional vector and each asset type as one dimension of that vector. By applying high-dimensional clustering, we divide all segments into several clusters based on their assets' conditions. The challenge in dealing with high-dimensional data lies in the "curse of dimensionality," a concept originally defined by Bellman (1961) that refers to the difficulty of optimizing a multi-variable function within the multi-dimensional context. In clustering, as dimensionality increases, the number of data points within each dimension becomes increasingly "sparse" (Steinbach et al., 2004). As illustrated in Figure 2.1, a dataset with 10 points is

randomly distributed from 0 to 1 in a one-dimensional space. The points are quite close to each other. There are four points within the range [0, 0.5]. But when the dataset is expanded to two-dimensional, if we still use 0.5 as the discretization unit in each dimension, there are then only 3 points in the range of [0, 0.5] in each dimension. When we further expand the dataset to three-dimensional, there are only 2 points within the same unit. So for high-dimensional data, distance may no longer be effective to distinguish points, and most cluster techniques applicable to low-dimensional data (e.g., centroid-based clustering, density-based clustering) are rendered meaningless.



(a) **(b)** **(c)**

Figure 2.1 Illustration of sparsely distributed data points due to curse of dimensionality

During the past decades, much effort has been devoted to avoid the curse of dimensionality. One solution is to develop new measurements for distance or similarity across clusters, including grid (Hinneburg and Keim, 1999), sum of similarities along dimensions (Aggarwal, 2001), and approximate similarity (Li et al., 2002). Li et al., 2002, suggested a practical similarity measurement called locality-sensitive hashing (LSH), which is a widely used algorithm to search for similarity between high-dimensional data for fast indexing and database searching. LSH maps high-dimensional data points to a low-dimensional space by applying hash functions. As mentioned by Datar et al. (2004), an eligible hash function family $H = \{h_1, h_2, h_3, \dots, h_i, \dots\}$ in LSH must be (d_1, d_2, p_1, p_2)-sensitive, which means for any two high-dimensional vectors q and v:

- if $\mathrm{D}(q, v) \leq \mathrm{d}_1$, then $P_H[\mathrm{h_i}(q) = \mathrm{h_i}(v)] \geq \mathrm{p}_1$
- if $\mathrm{D}(q, v) > \mathrm{d}_2$, then $P_H[h_i(q) = h_i(v)] \leq p_2$

where d_1 and d_2 are the critical distances to determine if q and v are similar, p_1 and p_2 are the critical probabilities, and D is distance measurement in the low-dimensional space. If the distance between the mapped values is less than or equal to d_1, then the probability that q and v are similar is greater than p_1. On the contrary, if the distance between mapped values is greater than d_2, then the probability of q and v being similar is less than p_2. Based on such definition, researchers proposed different function schemes and validated their reliability in capturing the underlying similarity, including inner product (Charikar, 2002), learned Mahalanobis distance (Jain et al., 2008), and normalized kernel function (Kulis and Grauman, 2012).

3. RESEARCH METHOD

Our proposed HDCSS method, as illustrated in Figure 3.1, consists of three major components: current condition estimation, similarity matrix construction, and stratification. Current condition estimation "predicts" asset conditions (e.g., in the form of LOM) based on historical records. This is to ensure that for the next round of inspection, sampling is conducted based on the former inspection results and the asset's deterioration rate. A similarity matrix containing the similarity between every two segments is constructed and applied to stratified sampling. Stratification is then accomplished via spectral clustering, which divides segments into strata (clusters). Segments within each stratum share similar patterns with regard to asset conditions. Thus, by selecting segments across strata, we select representative samples across all patterns. The sample size is a fixed percentage of segments in the network, constrained by labor or budget limits. The same percentage of segments within each stratum are chosen randomly. Once the sampled segments are inspected, maintenance and rehabilitation (M&R) activities can be further conducted on those segments whose performances are below a certain threshold. The M&R records and inspection results will then be applied to the next round of sampling process for inspection.

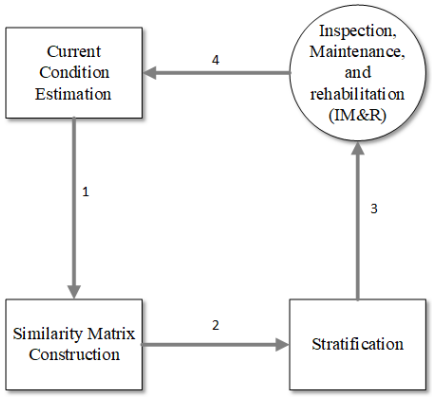


Figure 3.1 Illustration of stratification in the HDCSS method

3.1 Current Condition Estimation

As the sampling unit for maintenance activities, roadway segment possesses multiple features, such as asset types (shoulder work, litter, weeds, sweeping, maintaining inlets, etc.), geometric characteristics (number of lanes, segment length, etc.), and traffic information (AADT, peak hour volume, etc.). Each segment can therefore be described as a high-dimensional vector:

$$S_n = \{a_{shoulder\ work}, a_{litter\ pickup}, a_{weed}, \dots;\ g_{length}, g_{lane\_num}, \dots; t_{AADT}, t_{Peak\_Vol}, \dots\}$$

where S_n refers to segment n, $1 \le n \le N$, N is the number of segments in the network, and a, g, and t refer to the features associated with asset, geometric, and traffic, separately. In this paper, we only consider asset conditions as segment features.

The current condition estimation starts with translating the asset deterioration process into a deterioration matrix. The asset conditions are described using 15 letter scores from A+ to F-, based on the percentage of deficient assets within a segment. A+ represents the best condition and F- the worst. In previous studies, asset deterioration has been treated as either a linear (Brint, 2006) or non-linear (Prozzi and Hong, 2008; Saha et al., 2014) process. Since the deterioration of roadway assets is a relatively slow process, and inspection is often conducted quite frequently (every three or six months), we assume that asset deterioration is a linear process; yet, the rate may vary across different asset types and/or different

segments. For example, on Segment *1*, the time during which the segment's shoulder condition deteriorates from A to A- is the same as the time from A- to B+. Yet in the meantime, the condition of littering might deteriorate from A to C. And on Segment *2*, while the shoulder deteriorates from A to A- on Segment *1*, the shoulder condition might deteriorate from A to B+.

The deterioration matrix is computed based on paired consecutive inspection records without any intervention (e.g., M&R) in between. In this study, we filtered out all consecutive inspection records whose latter result outperformed the former one. Yet exceptions might occur when the asset might still deteriorate to a worse condition even after repair or maintenance, in which case this method might underestimate the asset deterioration rate.

To simplify calculation, the 15 letter grades of assets from A+ to F- are converted to numerical scores of 15 to 1. The deterioration process thus is equivalent to the score decreasing over time. The deterioration rate of any asset within a segment is calculated as the score difference divided by time duration between two inspections.

The historical records in our study span many years, during which segments may be inspected multiple times. Therefore, some segments might have more than one pair of consecutive records, producing different historical deterioration rates. In such cases, the average of historical deterioration rates is employed as the deteriorate rate of the segment. For segments without such prior records, deterioration rate is replaced with the network-averaged value. For example, if no consecutive record for shoulder work is available on one segment, its deterioration rate of that segment is replaced with average shoulder work deterioration rate of all segments. Deterioration matrix is constructed as:

$$D = \begin{bmatrix} d_{seg1\_ShoulderWork}, & d_{seg1\_LitterPickup,} & d_{seg1\_IceSnow}, & \cdots \\ d_{seg2\_ShoulderWork}, & d_{seg2\_LitterPickup}, & d_{seg2\_IceSnow}, & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ d_{segN\_ShoulderWork}, & d_{segN\_LitrerPickup}, & d_{segN\_IceSnow}, & \cdots \end{bmatrix} \quad (1)$$

D can always be updated with the latest inspection results and maintenance activities.

With the deterioration matrix constructed, we can estimate current conditions of assets on each segment. Previous asset conditions within the entire network can be expressed as:

$$M_{Previous} = (S_1, S_2, \ldots, S_N)^T \quad (2)$$

where *M* represents the previous network asset inspection conditions, *S* represents the asset conditions within the segment from previous inspection, with:

$$S_i = (s_{i\_ShoulderWork}, s_{i\_LitterPickup}, s_{i\_IceSnow}, \ldots) \quad (3)$$

The current asset conditions are then estimated by considering previous conditions and inspection frequency, which is expressed as:

$$M_{Current} = M_{Previous} + tD \quad (4)$$

where $M_{Current}$ is the estimated current asset conditions within the entire network, and t is time duration between previous and current inspections.

3.2 Similarity Matrix Construction

With the current asset conditions estimated, LSH is implemented to define the similarity between segments. All segments are then divided into clusters based on the similarity matrix via spectral clustering. A fixed percentage of segments can then be randomly chosen from each cluster. Figure 3.2 illustrates the LSH process in detail.

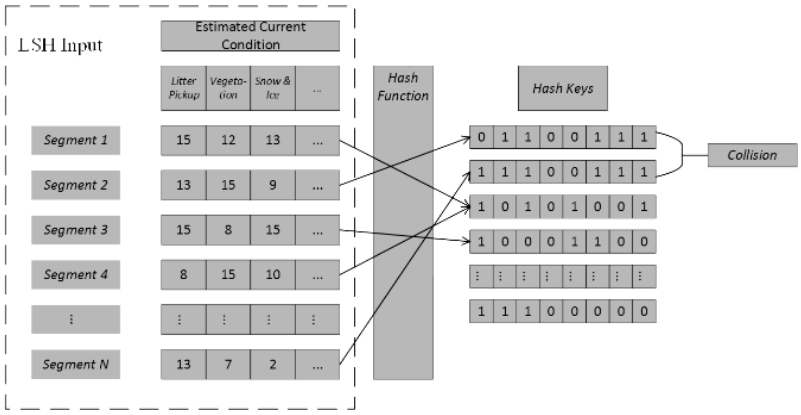


Figure 3.2 Illustration of LSH process

The input to LSH algorithm is the estimated current asset conditions, including:

$$M_{current} = (S_1^*, S_2^*, \dots, S_N^*)^T \tag{5}$$

$$\mathrm{S_i^*} = \{s_{i\_ShoulderWork}^*, s_{i\_\ LitterPickup}^*, s_{i\_IceSnow}^*, \cdots\} \tag{6}$$

where S_i^* represents the estimated current asset conditions on Segment i.

The first step in LSH is to define hash functions. In this study, we use inner product hash functions proposed by Charikar (2002). Hash function transforms a k-dimensional segment into a binary string. For example, in Figure 3.2, each segment is transformed into an 8-digit binary string. To determine the first digit of a binary string, we pick a k-dimensional vector $\boldsymbol{r} = (r_1, r_2, r_3, \dots, r_k)$. Each dimension ($r_1$,$r_2$, …) in vector $\boldsymbol{r}$ is randomly generated following Gaussian distribution. Then we calculate the inner product between the segment and r as:

$$h = \boldsymbol{r} \cdot S^* = r_1 s_{ShoulderWork}^* + r_2 s_{LitterPickup}^* + r_3 s_{IceSnow}^* + \cdots \tag{7}$$

where h is the inner product. When h is greater than or equal to 0, the first digit of the binary string is 1, and 0 otherwise. By repeating the process eight times, an 8-digit binary string is generated. The binary strings are the hash keys of a hash function family. The same process applies to all segments, with each segment assigned a hash key. Note that some segments may have the same hash keys.

Since hash keys are binary strings, we use Hamming distance as the difference measurement to compare them (Kulis and Grauman, 2009). For any two strings with equal length, Hamming distance is defined as the number of different digits. As illustrated in Figure 3.3, Hamming distance between the two strings is 1.

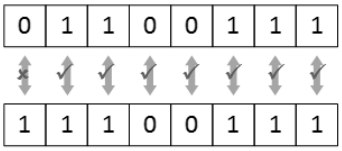


Figure 3.3 Illustration of Hamming distance

When the difference between two segments' hash keys is less than a certain threshold, it is referred to as "collision" between the two, and they are deemed similar. As illustrated in Figure 3.2, hash key of Segment *1* is 01100111, and hash key of Segment *N* is 11100111. Hamming distance between the two hash keys is 1. If the threshold to define a collision is set as 2, the difference between Segments *1* and *N*'s hash keys meets the requirement, and thus the two segments are deemed similar.

Until this step, LSH algorithm is fully implemented. Yet the algorithm can only determine whether two segments are similar or not rather than quantifying such similarity. Considering that hash function utilizes randomly generated vectors, using different vectors would lead to different hash keys. In Figure 3.2's example, Segment 1 and Segment *N* are considered similar. But if we generate another eight vectors, there is a probability that Segment 1 and Segment *N* are no longer similar. To remedy this, we perform LSH algorithm multiple times (i.e., 300 runs), and define similarity as the probability that two segments are similar across all runs. For example, if Segment 1 and Segment *N* are identified as similar for 240 of 300 times, the similarity between them is $240/300 = 0.8$. With the similarity between each pair of segments in the network quantified, a matrix of similarity $\mathbb{S} = [\mathrm{Sim}_{ij}]$is constructed, where Sim_{ij} represents the similarity between segments i and j.

3.3 Stratification

Once the similarity matrix of segments is constructed, we apply spectral clustering for stratification, which is one of the most popular clustering algorithms due to its simplicity and efficiency (Von Luxburg, 2007). It is originated from partitioning clustering, which gives weights to links between data points and divides clusters by removing the least weighted links between clusters. Spectral clustering combines partitioning clustering with graph Laplacian matrices. The calculation is based on the spectrum of similarity matrix. The detailed computation is available in Appendix I. By applying spectral clustering, the network is divided into several strata, and segments with similar asset conditions are classified into the same stratum.

4. APPLICATION

4.1 Study Site and Data

We implemented the HDCSS method using roadway asset inspection records provided by the Utah Maintenance Management Quality Assurance Plus (MMQA+) Program. Previously, MMQA+ performed full inventory inspection for asset maintenance. The maintenance personnel recorded total number of assets to be maintained and deficient assets on each segment. Then inspection records were entered into MMQA+ software to calculate the LOM (letter grade). In Utah, the entire highway network is divided into 489 segments. Inspection was performed semi-annually from September 2014 to March 2016, with several segments inspected multiple times within one inspection period. The inspection record archives overall asset condition (A+, A, A-, …), as well as segment ID, asset type, inspection date, and deficiency locations. With more than 7,000 records in the database, 14 asset types are used in our study, including shoulder work (SW), curb & gutter (CG), litter pickup (LP), weed control (WC), grade & clean ditches (GCD), maintain inlets (MI), erosion repair (ER), pavement markings (PM), repair & replace signs (RRS), repair & replace delineation (RRD), guardrail maintenance (GM), sweeping (SP), vegetation control (VC), and fence maintenance (FM).

Table 4.1 shows the average deterioration rate for each asset. For example, the average deterioration rate of CG is 0.0996. It means that the condition of CG deteriorates by 1 level (from A+ to A, or from A to A-) in approximately 10 months on average. Yet on individual segments, the rate can be different. For example, deteriorate rate of CG on some segment can be 0.2, indicating that it takes five months for CG to deteriorate from A+ to A.

Table 4.1 Average Deterioration Rates of Assets (per Month)

Asset	SW	CG	LP	WC	GCD	MI	ER
Rate	0.0756	0.0996	0.0821	0.0151	0.0394	0.0698	0.0813
Asset	PM	RRS	RRD	GM	SP	VC	FM
Rate	0.0181	0.0988	0.0244	0.0752	0.0022	0.0207	0.0825

In HDCSS, we use a 14-digit binary string as the hash key. The "collision" threshold is set as 2, indicating that when Hamming distance between the hash keys of two segments is less than 2, those two segments are similar. This number is determined by assessing the trade-off between number of similar segments and the similarity value. To avoid too many or too few segments in each cluster, all segments were divided into 10 clusters. For comparison purposes, SRS is also conducted on the same dataset. In the following section, both methods have been performed 50 times for sensitivity analysis.

4.2 Results and Analysis

The purpose of asset inspection is to evaluate asset conditions and report LOMs of overall highway network for investment decisions. Ideally, LOM distribution, measured from samples, can reflect both overall condition and condition variation. To assess the effectiveness of our sampling method, the difference between conditions estimated by samples and full inventory is computed with Root-Mean-Squared-Error (RMSE). For any asset, LOM distribution is expressed as $(X_{A+}, X_A, X_{A-}, \dots, X_{F-})$, where X_i is the actual percentage of grade i in the full inventory (all segments). The grade distribution estimated from sample is expressed as $(x_{A+}, x_A, x_{A-}, \dots, x_{F-})$, where x_i is the estimated percentage of grade i among all the sampled segments.

The RMSE between estimated (from sample) and ground-truth LOM distributions is then calculated as:

$$\text{RMSE} = \sqrt{\frac{\Sigma\,(x_i - X_i)^2}{15}} \tag{8}$$

RMSE reflects the error induced during the sampling process. As the value increases, estimated condition deviates from the ground-truth. To compare the performance of HDCSS with SRS methods, we consider the most recent inspection record of each segment as the ground-truth to compare against; and the second most recent inspection record as the historical record based on which to estimate the "current" asset conditions. To validate the robustness of sampling methods, particularly their sensitivity to different data dimensionalities, a series of sensitivity tests are performed. Using the same highway network, sampling is conducted with 6, 8, 10, and 14 different types of asset, separately. The types of assets are randomly selected when they are less than 14. Figure 4.1 shows the average RMSEs when the HDCSS method is applied based on sample rates ranging from 5% to 30% of the entire segment inventory. The sampling rate for the HDCSS method refers to the percentage of samples chosen from each stratum. Since the number of samples in each stratum is rounded to the nearest integer, the number of samples varies slightly from time to time. For example, when the sample rate is 10% in a network with 489 segments, the total number of segments chosen from each stratum is not always exactly 49. To justify its equivalence of sample sizes, the number of samples in the SRS method, which is also segment-based, is the same as the number of segments chosen by the HDCSS method. It is noted from Figure 4.1, for low-dimensional (e.g., less than 10 asset types) data, the average RMSEs show no significant difference when dimensionality changes. Yet when the dataset becomes high-dimensional (more than 10 asset types), the average RMSEs start to demonstrate improvements. It further validates the effectiveness and suitability of the HDCSS method for high-dimensional data. The LSH algorithm is designed for high-dimensional space where Euclidean distance is no longer valid as similarity measurements. As dimensionality increases, the HDCSS method tends to provide more accurate LOM estimation of the overall asset condition.

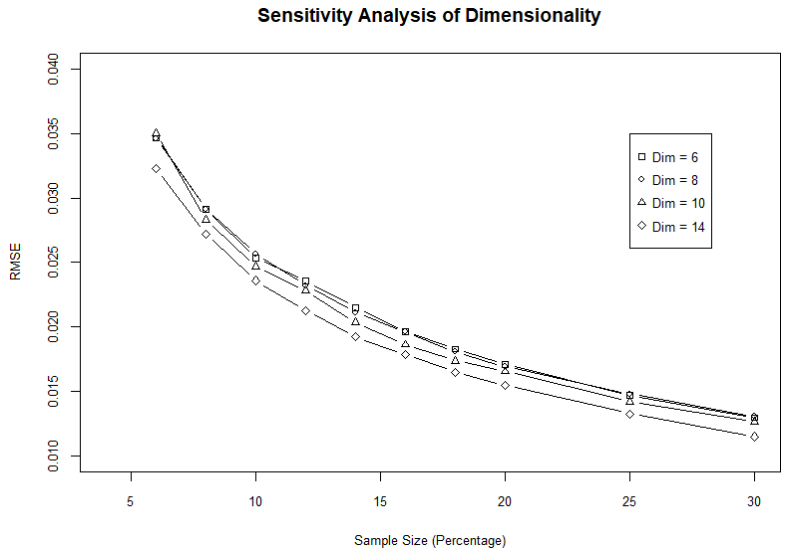


Figure 4.1 Sensitivity analysis of dimensionality (number of asset types) with different sample sizes (HDCSS)

Figure 4.2 (a) shows the sensitivity analysis of varying sample sizes for HDCSS versus SRS. Note that the RMSE is averaged out both across LOMs and across assets. As shown in Figure 4.2 (a), there is a trade-off between accuracy and sampling rate. Both average RMSE and its standard deviation decrease as the sampling rate increases. The slope appears to be steeper when the sampling rate is below 15%. The HDCSS method constantly outperforms SRS by providing lower average RMSE. Figure 4.2 (b through d) shows the RMSE distributions of both methods being applied for 200 times when sample sizes are 6%, 8%, and 10%, respectively. When sample size is less than 10%, there is a distinct difference between the performances of two sampling methods.

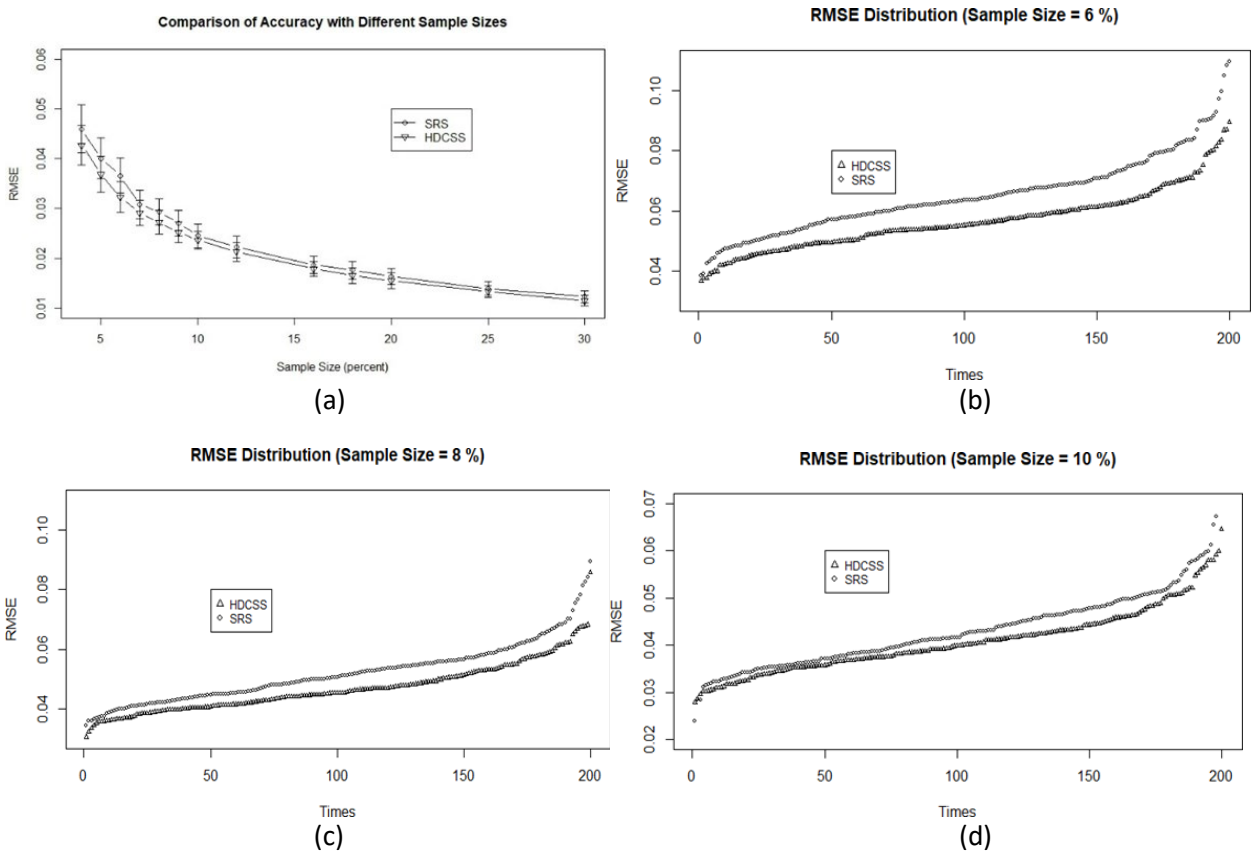


Figure 4.2 Sensitivity analysis of sample sizes between SRS and HDCSS methods

One highlight of HDCSS is that the selected sample segments can accurately reflect LOMs of different assets, respectively, throughout the network. Figure 4.3 provides a detailed look at the sampling accuracy for each asset type using SRS and HDCSS, where RMSE (mean and standard deviation) is shown for all 14 assets with a 20% sampling rate. As seen in Figure 4.3, the RMSEs vary significantly across assets. For most assets, SRS has higher RMSE than HDCSS, indicating the superiority of HDCSS. However, note that for certain assets (WC, RRD, SP, and VC), SRS yields lower RMSE. To further explore the underlying reasons, we compared the LOM distributions between each type of these assets.

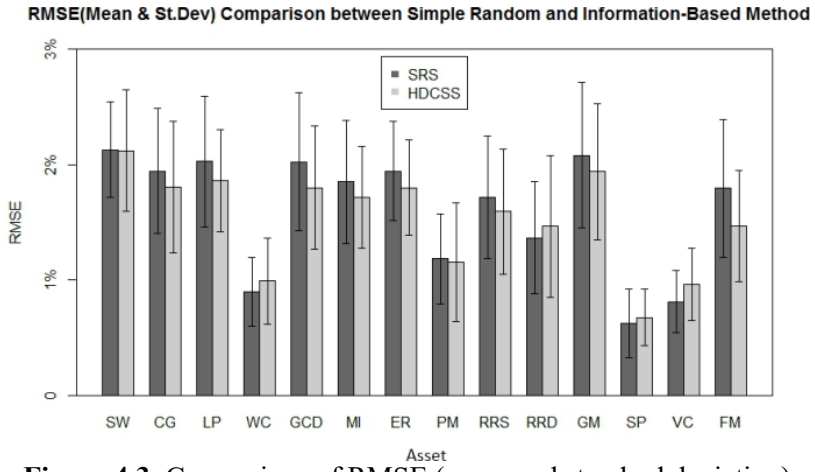


Figure 4.3 Comparison of RMSE (mean and standard deviation) between SRS method and HDCSS method

Figure 4.4 shows the ground-truth grade distributions of WC, RRD, SP, and VC. All these assets have a dominant grade (A+) and little variation across other grades, which explains the higher RMSE of HDCSS than SRS. For these assets, more than 80% of segments are of A+ grade. Under such circumstance, since the difference between individual samples is insignificant, it is highly likely that choosing different samples would not influence the result much. An extreme case in such a situation is that if all segments are of grade A+, then samples selected by any method would yield the same result. For assets with such skewed LOM distribution, both methods would estimate the overall conditions with low errors. Yet one unique aspect of HDCSS is when one dimension in the high dimensional vectors lacks variation, clustering relies more on other dimensions, and the importance (weight) of that dimension thus diminishes. Correspondingly, the overall condition of that asset (with little variation) is less represented in the samples selected by HDCSS than a random pick. That explains the underlying reason of low RMSEs for these four asset types and the outperformance of SRS. The estimation accuracy thus can be inferred from the LOM distribution. The explanation can also be verified by the RMSE ranking across the four assets. As shown in Figure 4.4, the four assets, ranked by the percentage of grade A+ for each type in descending order, are SP, VC, WC, and RRD. This order is exactly the same as the RMSE ranking using both methods.

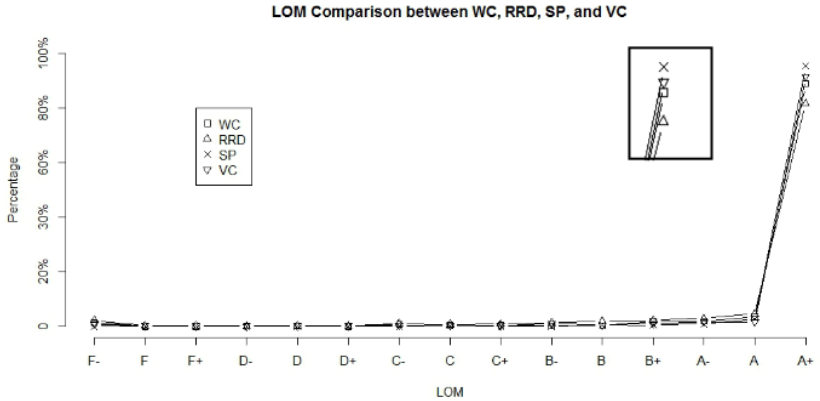


Figure 4.4 Grade distribution (ground truth) comparison between assets: WC, RRD, SP, and VC

In asset inspection sampling, one prominent concern is to reduce the sampling rate, since it ties directly to costs and budget allocation. According to Schmitt et al., (2006), some DOTs assume that asset conditions follow distributions, so a sampling rate can be estimated with a given confidence interval and accuracy. For example, North Carolina DOT performs sampling based on 90% to 95% confidence interval and 6% accuracy. Virginia DOT requires the confidence interval of sampling to be at 95% and accuracy of 4%. Since this study is not based on such assumptions, the comparison between sampling results and ground-truth is a non-parametric problem. Therefore, we use two methods to verify the sampling results of HDCSS.

The first method is to define a new index for measuring the accuracy of sampling results. We define an "accuracy rate" for each method, representing the probability of a sample being considered as accurate within a certain error threshold. The sampling result is considered "accurate" only if the errors between estimated asset conditions and ground-truth are within an acceptable range. The error can be quantified via RMSE. For example, if RMSE threshold is 0.05, the sampling result is not considered accurate unless the RMSEs of all assets are less than or equal to 0.05. If the sampling results of a method with a certain sample size are accurate for 80 of 100 times, the accuracy rate is 80/100 = 0.8. Figure 4.5 shows the sensitivity analysis of accuracy rate when different sample sizes apply. It is noted that under the same error threshold, when the sampling rate is less than 20%, HDCSS always yields a higher accuracy rate than SRS. For an accuracy rate of 85% (170 of 200 times) with an error threshold of 0.07 (marked as A), the required sampling rate is around 8% for HDCSS as opposed to 10% for SRS. To achieve an accuracy rate of 95% (190 of 200 times) with an error threshold of 0.05 (marked as B), by using the HDCSS method, the sample size can be reduced from 20% to 16%. Such decreased sample size reduces inspection costs for asset management, especially for large-scale highway networks.

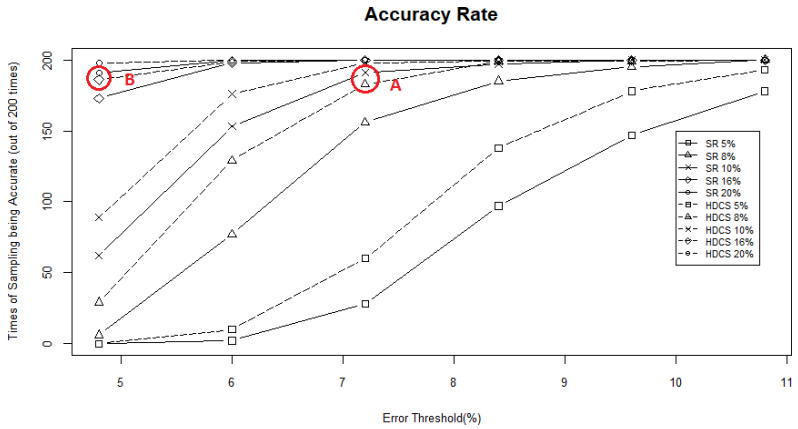


Figure 4.5 Sensitivity analysis of accuracy rates under different error thresholds and sample sizes

Another method is the ANOVA test to compare sampling results from different methods. To further explore sampling rate reduction quantitatively, we performed a one-way ANOVA test to analyze the difference of errors between the two sampling methods with varying sampling rates. Table 4.2 shows the result of the ANOVA test between the errors of samples selected by the HDCSS method with a 16% sampling rate and the SRS method with an 18% sampling rate. It shows there is no significant difference between the errors of estimated asset conditions. That is to say, for any ongoing sampling scheme using the SRS method with 18% of the entire population as a sampling rate, our method can effectively reduce it to 16%.

Table 4.2 Result of ANOVA for errors of samples selected by HDCSS (16%) and SRS (18%) methods ($\alpha = 0.05$)

	Degree of Freedom	Sum of Squares	Mean Square	F-Value	P-Value	Significantly Different, Yes/No
Between Features	1	0.0000033	3.314e-06	1.28	0.259	No
Within Features	398	0.0010306	2.589e-06			

Similar sample rate reduction results have been observed under other precision requirements, as shown in Table 4.3. It is observed that when the sampling rate of SRS is below 15%, HDCSS can reduce the sample rate by 1%. When the sampling rate of SRS is above 15%, HDCSS can reduce the sampling rate by 2%. If we assume the inspection costs are proportional to the number of segments to inspect, asset management agencies can reduce the sample rate from 9% to 8% by using HDCSS, saving 11% ([9%-8%]/9% = 11%) on their inspection budgets. Based on the sample rate reduction in Table 4.3, the reduced sample rate can save inspection costs by 7% to 11%. Such an effect can become more pronounced when applied to large highway networks.

Table 4.3 Results of ANOVA tests ($\alpha = 0.05$)

HDCSS Sample Rate (%)	SRS Sample Rate (%)	P-Value
8	9	0.51
10	11	0.73
11	12	0.806
12	13	0.814
16	18	0.119
18	20	0.259
19	21	0.298

By comparing the RMSEs of HDCSS and SRS, we can conclude that HDCSS outperforms SRS when applied to Utah's highway network. To further test the performance of HDCSS with data from different LOM distributions and evaluate the effectiveness of HDCSS under various scenarios, we have created a new high-dimensional dataset with different LOM distributions. The new dataset models a network with 500 segments and 20 types of assets on each segment. The LOMs of assets in the network, ranging from A+ to F-, are converted into numerical scores from 15 to 1. The initial asset LOMs are generated with Monte Carlo simulation. There are two distributions the initial asset conditions follow. The initial scores X for 10 randomly chosen asset types follow normal distribution $\mathrm{X}\sim\mathrm{N}(\mu,\sigma)$, where $4 \leq \mu \leq 13$ and $1 \leq \sigma \leq 3$. The initial scores Y of the remaining 10 asset types are generated with Poisson distribution, where $\mathrm{Y} = 16 - \mathrm{Z}$, $\mathrm{Z}\sim\mathrm{Pois}(\lambda)$, and $1 \leq \lambda \leq 5$. Note that we assume inspection to be conducted every six months. To model the deterioration process, we assume that deterioration between two consecutive inspections follows exponential distribution $Exp(\lambda)$, where $0.5 \leq \lambda \leq 1$. The values of parameters in all aforementioned distributions are empirically determined. We simulated the deterioration process for a two-and-a-half-year period, totaling six inspections. The asset conditions from the first five inspections are considered as the historical maintenance records. The asset conditions from the sixth inspection are considered as the ground-truth conditions. Two consumptions are made in terms of maintenance and inspection activities:

1. The first five inspections are full inventory inspections.
2. Any assets identified in condition F- will be repaired immediately. These assets' conditions therefore return back to A+ for the next inspection.

With this simulated dataset, we repeated the sampling procedure with both HDCSS and SRS methods for 200 times and analyzed sampling results. The comparison of RMSEs between the two methods is shown in Figure 4.6. It is clear that SRS yields higher RMSE compared with our HDCSS method.

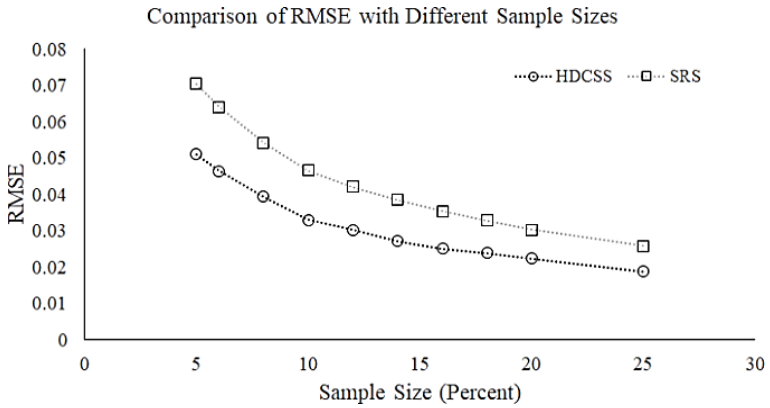


Figure 4.6 Comparison of RMSE between HDCSS and SRS

5. CONCLUSIONS

An HDCSS method is proposed. The sampling segments selected by this method can accurately represent the overall asset conditions of the full inventory. The HDCSS method generally outperforms the SRS method, which is widely used by DOTs. The method consists of three components: current condition estimation, similarity matrix construction, and stratification. The current condition estimation aims to provide predicted asset condition for clustering based on historical inspection records. Segments with multiple asset types are treated as high-dimensional vectors. By applying LSH algorithm and spectral clustering, the similarity between segments is measured and the segments are assigned to clusters (strata). Using inspection records from the State of Utah, it is verified that HDCSS outperforms SRS for most asset types, especially under the circumstances when LOM varies greatly within assets. For the assets with few LOM variations across segments, both the HDCSS method and SRS yield low errors. Our proposed method requires a relatively smaller sample size, leading to a potential decrease in inspection costs, especially for large-scale networks.

By using the HDCSS method, DOTs can save resources and time for asset inspection, because inspection is carried out on the segment basis, and the similarity identification introduced through LSH algorithm. The method can be further applied to any high-dimensional sampling process, e.g., in selecting corridor segments, intersections, or traffic assets where multiple types of features, e.g., traffic condition, geometric design, or assets, need to be considered. Based on this study, two intriguing topics emerge. First, as an important component of the method, deterioration matrix construction can significantly influence the accuracy of the sampling method. It is necessary to apply a more rigorous data analysis tool to enhance the estimation of the deterioration process. Second, it might be interesting to involve other, more efficient, HDC methods in the sampling process, which can potentially improve the accuracy of sampling results.

REFERENCES

Adams, T.M., Winkelman, B., 2016. "Statistical Analysis for Assessing Highway Maintenance Level of Service." *Transp. Res. Rec. J. Transp. Res. Board 2551*, 73–81. doi:10.3141/2551-09

Aggarwal, C., 2001. "Re-Designing Distance Functions and Distance-Based Applications for High Dimensional Data." *ACM SIGMOD Rec. 30*, 256–266. doi:10.1145/375551.383213

Bellman, R.E., 1961. *Adaptive Control Processes: A Guided Tour.*

Ben-Akiva, M., Humplick, F., Madanat, S., Ramaswamy, R., 1993. "Infrastructure Management under Uncertainty: Latent Performance Approach." *J. Transp. Eng. 119*, 43–58. doi:10.1061/(ASCE)0733-947X(1993)119:1(43)

Brint, A.T.T., 2006. "Sampling on Successive Occasions to Re-Estimate Future Asset Management Expenditure." *Eur. J. Oper. Res. 175*, 1210–1223. doi:10.1016/j.ejor.2005.06.031

Charikar, M.S., 2002. "Similarity Estimation Techniques from Rounding Algorithms." Proc. Thiry-Fourth Annu. ACM Symp. Theory Comput. - STOC '02 380–388. doi:10.1145/509907.509965

Datar, M., Immorlica, N., Indyk, P., Mirrokni, V.S., 2004. "Locality-Sensitive Hashing Scheme Based on P-Stable Distributions." Proc. Twent. Annu. Symp. Comput. Geom. 253–262. doi:10.1145/997817.997857

De la Garza, J.M., Piñero, J.C., Ozbek, M.E., 2008. "Sampling Procedure for Performance-Based Road Maintenance Evaluations." *Transp. Res. Rec. J. Transp. Res. Board 2044*, 11–18. doi:10.3141/2044-02

Durango-Cohen, P., Madanat, S., 2008. "Optimization of Inspection and Maintenance Decisions for Infrastructure Facilities under Performance Model Uncertainty: A Quasi-Bayes Approach." *Transp. Res. Part A Policy Pract. 42*, 1074–1085. doi:10.1016/j.tra.2008.03.004

Durango-Cohen, P.L., 2007. "A Time Series Analysis Framework for Transportation Infrastructure Management." *Transp. Res. Part B Methodol. 41*, 493–505. doi:10.1016/j.trb.2006.08.002

Guillaumot, V.M., Durango-Cohen, P.L., Madanat, S.M., 2002. "Adaptive Optimization of Infrastructure Maintenance and Inspection Decisions under Performance Model Uncertainty." *J. Infrastruct. Syst. 9*, 133–139.

Hinneburg, A., Keim, D.A., 1999. "Optimal Grid-Clustering: Towards Breaking the Curse of Dimensionality in High-Dimensional Clustering," in: Proceedings of the 25th International Conference on Very Large Databases. pp. 506–517.

Jain, P., Kulis, B., Grauman, K., 2008. "Fast Image Search for Learned Metrics," in: Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR). pp. 1–8. doi:10.1109/CVPR.2008.4587841

Kulis, B., Grauman, K., 2012. "Kernelized Locality-Sensitive Hashing." *IEEE Trans. Pattern Anal. Mach. Intell. 34*, 1092–1104. doi:10.1109/TPAMI.2011.219

Kulis, B., Grauman, K., 2009. “Kernelized Locality-Sensitive Hashing for Scalable Image Search.” *Proc. IEEE Int. Conf. Comput.* Vis. 2130–2137. doi:10.1109/ICCV.2009.5459466

Li, C., Chang, E., Garcia-Molina, H., Wiederhold, G., 2002. “Clustering for Approximate Similarity Search High-Dimensional Spaces.” *IEEE Trans. Knowl. Data Eng.* 14, 792–808. doi:10.1109/TKDE.2002.1019214

Liu, X.C., Chen, Z., 2017. “Spatial Sampling with Fisher Information for Optimal Maintenance Management and Quality Assurance.” *J. Transp. Eng.* Part A Syst. 143, 1–9. doi:10.1061/JTEPBS.0000083.

Medina, R. a., Haghani, A., Harris, N., 2009. “Sampling Protocol for Condition Assessment of Selected Assets.” *J. Transp. Eng. 135*, 183–196. doi:10.1061/(ASCE)0733-947X(2009)135:4(183)

Memarzadeh, M., Pozzi, M., 2016. “Integrated Inspection Scheduling and Maintenance Planning for Infrastructure Systems.” *Comput. Civ. Infrastruct. Eng. 31*, 403–415. doi:10.1111/mice.12178

Mishalani, R.G., Gong, L., 2009a. “Optimal Infrastructure Condition Sampling over Space and Time for Maintenance Decision-Making under Uncertainty.” *Transp. Res. Part B Methodol. 43*, 311–324. doi:10.1016/j.trb.2008.07.003

Mishalani, R.G., Gong, L., 2009b. “Evaluating Impact of Pavement Condition Sampling Advances on Life-Cycle Management.” *Transp. Res. Rec. J. Transp. Res. Board 2068*, 3–9. doi:10.3141/2068-01

Mishalani, R.G., Gong, L. 2009c. “Optimal Sampling of Infrastructure Condition: Motivation, Formulation, and Evaluation.” *J. Infrastruct. Syst. 15*, 313–320. doi:10.1061/(ASCE)1076-0342(2009)15:4(313)

Ng, A.Y., Jordan, M.I., Weiss, Y., 2001. “On Spectral Clustering: Analysis and an Algorithm.” *Adv. Neural Inf. Process. Syst. 14* 14, 849–856. doi:10.1.1.19.8100

Prozzi, J.A., Hong, F., 2008. “Transportation Infrastructure Performance Modeling through Seemingly Unrelated Regression Systems.” *J. Infrastruct. Syst. 14*, 129–137.

Saha, P., Liu, R., Melson, C., Boyles, S.D., 2014. “Network Model for Rural Roadway Tolling with Pavement Deterioration and Repair.” *Comput. Civ. Infrastruct. Eng. 29*, 315–329. doi:10.1111/mice.12057

Schmitt, R.L., Owusu-ababio, S., Weed, R.M., Nordheim, E. V, Schmitt, R.L., 2006. “Development of a Guide to Statistics for Maintenance Quality Assurance Programs in Transportation.”

Smilowitz, K., Madanat, S., 2000. “Optimal Inspection and Maintenance Policies for Infrastructure Networks.” *Comput. Civ. Infrastruct. Eng. 15*, 5–13. doi:10.3141/1667-01

Steinbach, M., Ertöz, L., Kumar, V., 2004. “The Challenges of Clustering High Dimensional Data,” in: New Directions in Statistical Physics. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 273–309. doi:10.1007/978-3-662-08968-2_16

Von Luxburg, U., 2007. “A Tutorial on Spectral Clustering.” *Stat. Comput. 17*, 395–416.

APPENDIX: SPECTRAL CLUSTERING ALGORITHM (NG ET AL., 2001):

Given a set of points $V = \{v_1, v_2, \dots, v_n\}$, the similarity matrix $S = \{s_{ij}\}$, where s_{ij} refers to the similarity between v_i and v_j, k number of clusters to construct:

1. Construct a fully connected similarity graph with S. Let D be its weighted adjacency matrix $D_{ii} = \Sigma_j S_{ij}$
2. Compute the normalized Laplacian $L = D^{-1/2} S D^{-1/2}$
3. Find $x_1, x_2, \dots, x_k$, the k largest eigenvectors of L, and form the matrix $X = [x_1, x_2, \dots x_k]$ by stacking the eigenvectors in columns
4. Form the matrix Y from X by renormalizing each of X's rows to have unit length (i.e. $Y_{ij} = X_{ij} / \left(\Sigma_j X_{ij}^2\right)^{1/2}$)
5. Treat each row of Y as a point in $\Re^k$, cluster them into k clusters via K-means
6. Finally, assign the original point v_i to cluster j if and only if row i of the matrix Y was assigned to cluster j