Automatic Real-time Detection and Correction of Erroneous Detector Data Using Fourier Transforms for On-line Traffic Control Architectures

(Paper #02-2244)

Srinivas Peeta\(^1\) and Ioannis Anastassopoulos\(^2\)

\(^1\)Srinivas Peeta
School of Civil Engineering
Purdue University
West Lafayette, IN 47907
Phone: (765) 494-2209
Fax: (765) 496-1105
E-mail: peeta@purdue.edu

\(^2\)Ioannis Anastassopoulos
School of Civil Engineering
Purdue University
West Lafayette, IN 47907
Phone: (765) 494-2206
Fax: (765) 496-1105
E-mail: anastass@purdue.edu

Accepted for publication in Transportation Research Record.

Manuscript Length: 6532
3 Figures: 750
2 Tables: 500
Total: 7782
ABSTRACT

Real-time and/or internet-based control architectures are currently being designed in the context of on-line route guidance and traveler information systems for vehicular traffic networks. They involve the transmission of field data to the Traffic Control Center (TCC) for real-time processing. To enable reliable and uninterrupted operation, these real-time systems should be fault tolerant to critical hardware failure modes such as malfunctioning detectors and failed transmission/communications links. This research proposes a Fourier transform-based fault tolerant framework to detect data faults due to malfunctioning detectors and predict the likely actual data for the seamless operation of an on-line traffic control architecture. However, incidents also exhibit data characteristics similar to those of some hardware related data faults. The proposed approach treats data faults and incidents as abnormalities in the monitored network. It first detects an abnormality and then distinguishes data faults from incidents, using a Fourier transform-based approach. Data faults are corrected using a Fourier transform-based data correction heuristic. Field data from Athens, Greece, and Hayward, California, are used to validate the proposed methodologies. The approach uses data directly without elaborate modeling, circumventing likely modeling errors and enabling simpler adaptability to future demand/supply changes. A key contribution is its ability to robustly predict near-term traffic conditions efficiently using historical data and the immediate past data on the current day. This alleviates the computational burden and enables minute-to-minute traffic prediction, substantially aiding the seamless and uninterrupted operation of on-line traffic control architectures.

Key words: Fourier transforms, fault tolerance, data correction, near-term traffic prediction
INTRODUCTION

The advent of intelligent transportation systems (ITS), fueled by advances in information technology, is motivating new paradigms for real-time traffic management in vehicular networks. Real-time traffic control architectures focus on operating traffic systems by measuring and processing time-dependent traffic flow data to develop strategies that enhance system performance. As documented in [1,2], detectors and/or communication links (between the field and the data processing center) can fail, raising key issues of fault tolerance that are characteristic of real-time systems. This study focuses on detecting data errors due to malfunctioning detectors and predicting their likely actual values using Fourier transforms-based techniques to enable the seamless operation of the associated real-time traffic control architecture. Addressing communication link failures is outside the scope of this research since they can be detected efficiently using simple computer communication techniques such as data stamping. Hence, the primary study objective is to ensure the continuous availability of reliable data for on-line traffic control architectures under detector failures.

An on-line control architecture for an automated remote real-time traffic system [3,4] that provides drivers real-time routing advisories, is used as the study context. As per this architecture, real-time traffic data from the field detectors is transmitted via communication links to the Traffic Control Center (TCC), and from there to a data processing center which may be located at another site. After the data is received at the processing center, it is used to generate routing strategies in sub-real time on a high-performance computing system. These strategies are transmitted to the TCC for dissemination to the drivers. The constituent elements of the associated solution procedures executed on the computing system include: (i) calibration and consistency checking, (ii) a virtual system simulator, and (iii) dynamic traffic assignment models. However, to ensure that the data used by these solution procedures is correct and uninterrupted, a fault tolerance component is used as a filter before the data is accessed by the solution procedures. This component, which forms the focus of the study, aims to develop methodologies to automatically detect detector malfunctions by continuously monitoring data and “correct” the associated faulty data by predicting the likely actual data. Such a mechanism ensures the seamless and uninterrupted real-time operation of the on-line traffic control architecture.

In a broader sense, the Fourier-based methodologies [5] developed as part of the fault tolerance component can be viewed as a mechanism to efficiently predict near-term future traffic conditions (in terms of speed, occupancy, volume) by using historical data and the immediate past data on the current day. Thereby, the approach can also be used to predict future travel times or automatically detect incidents. In view of space constraints here, the details of the incident detection component and its analysis are discussed elsewhere by the authors [5, 6]. The erroneous data due to detector malfunction is characterized here as a fault in the real-time traffic system and is labeled data fault. Hence, the objective of this paper is to automatically detect data faults and predict the likely actual data for the uninterrupted operation of an on-line traffic control architecture. Since incidents and data faults carry similar data signatures [5, 6], the proposed methodology first differentiates between them before correcting confirmed data faults.

Fault detection has been explored widely in various fields. In the Transportation arena, existing research primarily focuses on the detection of loop detector errors. It can be classified
as: (i) microscopic; where the detection procedure is performed by the detector’s microprocessor, and (ii) macroscopic; where it is undertaken at the processing center. Typically three criteria are used for error detection at the microscopic level. First, pulses or actuation gaps shorter than a pre-specified threshold value are ignored, accounting for small drop-outs in vehicle presence as well as false detector actuations. It can be viewed as a direct signal refinement procedure. Second, if more than two valid pulses are sensed during one second, they are flagged as unreliable. The third criterion compares the detector on-time and off-time against certain pre-programmed variables. The detector on-time is the time for which the detector senses the presence of vehicle, while the off-time is the opposite.

In a previous Federal Highway Administration (FHWA) study [7], five possible loop detector failures were identified: (i) stuck sensors (on or off), (ii) chattering, (iii) pulse break-up (pulsing), (iv) hanging (on or off), and (v) intermittent malfunctioning. Several algorithms have been proposed to address them. The “single-detector algorithm” compares detector data (volume \( V \), occupancy \( O \), and speed \( S \)) with threshold values at various time intervals [8]. This approach can identify many malfunctioning states, and if the comparisons are performed over short time intervals, it can also detect intermittent data errors.

A more recent microscopic approach uses the inherent advantages of dual loop detectors [9]. A vehicle passing over a dual loop detector generates two measurements, one from the upstream loop and one from the downstream loop. The algorithm is based on the assumption that since the two loops are placed close to each other speed is maintained, and therefore the on-time should be identical for the two loops. If not, an error is detected. However, there is a significant possibility for the two on-times to be different at lower speeds even under no malfunction, due to acceleration or deceleration. To address this issue, a statistical procedure is incorporated in the algorithm.

Microscopic approaches require the reprogramming and/or modification of the detector device, and primarily focus on inductive loop detectors. Recent developments in the ITS arena have spawned a new generation of sophisticated sensor devices such as microwave radar, ultrasonic, and laser radar detectors. Current microscopic approaches cannot be used for these sensor devices. A macroscopic approach is typically opaque to the sensor type and is based mostly on data processing. Hence, a macroscopic approach is preferred in our on-line architecture context.

Macroscopic approaches typically compare volumes and/or occupancies with built-in thresholds. If these thresholds are violated, an error is assumed. However, only some gross errors can be detected using such procedures. Another approach [10] compares a detector’s average on-time to that average across all detectors at the detector’s station. It detects hanging-on (detector jammed at on position) and hanging-off (detector jammed at off position) states but requires extensive calibration. Also, since its logic is based on the comparison across lanes, changes in freeway traffic flow behavior may decrease reliability.

An improved macroscopic-level approach [11] uses three different threshold tests: (i) a \( V/O \) threshold test for 20-second interval data, (ii) a volume threshold test, also for 20-second intervals, and (iii) a volume threshold test for 5-minute intervals. The decision leading to the detection of an error is based on a decision tree technique. This algorithm has a three-dimensional error-checking approach as speed is also used. Analysis indicates a robust ability to
detect intermittent failures and a low false alarm rate. However, it cannot detect hanging-on malfunctions and has a high dependency on threshold values.

Neural network models and genetic algorithms have also been used to address fault tolerance issues, mostly in other engineering domains. A detailed discussion on their application for fault diagnosis and correction is provided in [5, 12].

Forecasting techniques are useful in this study context to correct data faults and predict near-term future traffic conditions. Well-known approaches include: the nearest-neighbor algorithm, the historical average, neural network applications, and autoregressive integrated moving average models (ARIMA) [13]. A more recent approach combines ARIMA models with Kohonen maps [14]. Seasonal ARIMA and seasonal exponential smoothing models [15] can take into account the cyclical characteristics of traffic flow. Hence, they can predict flow more realistically, especially near peak periods. They are capable of 15-minute single interval predictions. Multiple-interval and shorter time interval predictions (of the order of 1-minute), which are necessary in the context of an on-line control architecture, are currently not addressed.

In summary, many fault detection approaches involve modeling, which has the potential for inherent errors. Others require extensive calibration and/or training. In the context of traffic systems, fault detection is addressed for loop detector errors. Many existing approaches are specific to loop detectors and cannot be generalized. Microscopic approaches involve modifications at the microprocessor level. Macroscopic approaches typically involve extensive calibration. Existing forecasting methods to determine traffic flow parameters contain several coefficients that require extensive calibration. Most are capable of 15-minute single interval predictions, which are not sufficient for the minute-to-minute monitoring requirements of many real-time traffic operation/control frameworks.

This research proposes a Fourier-based framework where data is analyzed directly, to detect and correct data faults. The approach is general and can be implemented for any sensor device that provides volume, speed, and occupancy. The fault correction approach is macroscopic and based on the same Fourier-based methodology used for its detection. It is capable of multiple-interval forecasts for the near- and medium-term, and is hence amenable to real-time traffic operations. It does not involve extensive calibration/training, enables straightforward computation of parameters, is computationally very efficient, and can adapt to demand/supply changes over time using a continuous learning feature.

**METHODOLOGY**

An approach to detect data faults requires a capability to distinguish between faults and incidents. This is because incidents have characteristics similar to some types of data faults. Incidents cause temporal traffic flow pattern changes in terms of an abrupt increase in the occupancy-time sequence, and sudden decreases in flow and speed upstream of the incident. Faults can also manifest as sudden decreases in speed or flow, or increased occupancy. However, these data changes are virtual vis-à-vis the traffic system and unlikely to simultaneously occur as in an incident. Since the study seeks to isolate data faults, problematic data is first identified as an abnormality, and then classified as a fault or an incident using the proposed Fourier-based methodology. If a data fault is identified, the likely correct data is predicted using a Fourier-based data correction heuristic.
Fourier Transforms

Fourier transforms can express any periodic function, or equivalently any function defined on a finite interval, exactly as an infinite sum of sinusoidal components [16]. If a function \( f(x) \) is defined for the interval \( 0 \leq x \leq c \), it can be represented by the Fourier series:

\[
f(x) = K + \sum_{n=1}^{\infty} A_n \cdot \cos\left(\frac{2n\pi x}{c}\right) + \sum_{n=1}^{\infty} B_n \cdot \sin\left(\frac{2n\pi x}{c}\right)
\]

(1)

where \( K \) is a constant, \( A_n \) and \( B_n \) are the Fourier coefficients of the cosine and sine components respectively, and \( c \) is time interval for which a Fourier transform is to be determined. All sinusoidal terms have angular frequencies (multiples of \( 2\pi/c \)) so that a whole number of wavelengths can be fit into \( 0 \leq x \leq c \). To enable tractability, (1) can be re-written as the finite sum of \( N \) sinusoidal components where a large enough \( N \) can ensure a sufficiently accurate approximation of \( f(x) \):

\[
f_N(x) = K + \sum_{n=1}^{N} A_n \cdot \cos\left(\frac{2n\pi x}{c}\right) + \sum_{n=1}^{N} B_n \cdot \sin\left(\frac{2n\pi x}{c}\right)
\]

(2)

The truncated discrete Fourier series \( f_N(x) \) is easier to compute and is hence useful for applications. A Fourier series has the following useful properties:

1) Multiplication Property

\[
\lambda \cdot f(x) = \lambda \cdot K + \sum_{n=1}^{\infty} (\lambda \cdot A_n) \cdot \cos\left(\frac{2n\pi x}{c}\right) + \sum_{n=1}^{\infty} (\lambda \cdot B_n) \cdot \sin\left(\frac{2n\pi x}{c}\right)
\]

(3)

2) Additive Property

If \( f(x) = K_1 + \sum_{n=1}^{\infty} A_n^1 \cdot \cos\left(\frac{2n\pi x}{c}\right) + \sum_{n=1}^{\infty} B_n^1 \cdot \sin\left(\frac{2n\pi x}{c}\right) \), and

\[
g(x) = K_2 + \sum_{n=1}^{\infty} A_n^2 \cdot \cos\left(\frac{2n\pi x}{c}\right) + \sum_{n=1}^{\infty} B_n^2 \cdot \sin\left(\frac{2n\pi x}{c}\right),\text{ then}
\]

\[
f(x) + g(x) = (K_1 + K_2) + \sum_{n=1}^{\infty} (A_n^1 + A_n^2) \cdot \cos\left(\frac{2n\pi x}{c}\right) + \sum_{n=1}^{\infty} (B_n^1 + B_n^2) \cdot \sin\left(\frac{2n\pi x}{c}\right)
\]

(4)

These properties also hold for the discrete case. In this paper the following complex algebra based expression of the Fourier series is used:

\[
f(x) = K + \sum_{n=1}^{N} \{ \text{Re}(n) \cdot \cos[2\pi \cdot \nu(n) \cdot x] + \text{Im}(n) \cdot \sin[2\pi \cdot \nu(n) \cdot x] \}
\]

(5)

where, \( \text{Re}(n) \) and \( \text{Im}(n) \) are the real and imaginary Fourier coefficients, respectively, and \( \nu(n) \) is the frequency. \( \nu(n) = \frac{n-1}{N \cdot \iota} \), where \( \iota \) is the time interval between two consecutive data points (in seconds), labeled data acquisition interval. The amplitude \( \text{ampl}(n) \) is equal to \( \sqrt{\text{Re}^2(n) + \text{Im}^2(n)} \).

Overview of the Fourier-based Methodology

The Fourier theory is used to analyze real-time streaming traffic data from field detectors to identify data faults and/or incidents. The commonly available traffic flow measures from
detectors include volume $V(t)$, speed $S(t)$, and occupancy $O(t)$, as a function of time $t$. Fourier transforms are used to infer on the underlying characteristics of this data to identify data abnormalities that occasionally arise. The main steps of the proposed methodology are: (1) initial data processing, (2) training, (3) detection of data abnormalities, (4) distinguishing between data faults and possible incidents, (5) data correction, and (6) incident detection. Step 6, which determines whether an abnormality identified as a possible incident is actually an incident, is not a focus of this paper and is discussed elsewhere [6].

Initial Data Processing
The historical abnormality-free data for a road cross-section is first smoothed to reduce the inherent volatility of real traffic data. Raw data is characterized by temporary traffic flow fluctuations representing noise in the acquired data. The smoothing is performed using an exponential smoothing technique. Suppose $d(t)$ is a data point for time $t$. Its smoothed data value $\tilde{d}(t)$ is obtained using $P$ preceding data points. Smoothing increases with increasing $P$. The exponential smoothing is obtained as follows:

$$\tilde{d}(t) = \frac{\sum_{i=1}^{P} e^{-\alpha} \cdot d(t - i + 1)}{\sum_{i=1}^{P} e^{-\alpha}}$$

(6)

where $\alpha$ is a smoothing parameter. Reducing $\alpha$ increases smoothing implying an increased contribution from past data points. For a small $\alpha$ ($\approx 0.01$), exponential smoothing is almost identical to moving average smoothing.

The smoothed historical abnormality-free data for time $t = t_0, \ldots, t_e$ is grouped into $N$-point time sequences. Here $t_0$ and $t_e$ are the first and last data points, respectively, in the historical data set. The first time sequence starts at $t_0$ and ends at $t_0 + (N-1)$. The next one is obtained by shifting the current sequence by one (starts at $t_0 + 1$ and ends at $t_0 + N$). Each data set contains $N$ data points which is equal to the number of Fourier points in order to minimize the approximation error of the transformation of the time sequence to its Fourier series. Using the maximum possible number of Fourier points minimizes this approximation; hence the number of Fourier points is also $N$ (since, for any data set containing $Z$ data points, the maximum number of Fourier points that can be used is $Z$). To speed up the computational process, the Fourier transforms are computed using a Fast Fourier Transform (FFT) algorithm that requires $N$ to be a power of 2 for strictly computational reasons.

Training
The first step of the training is performed by computing the Fourier transform and the Total Fourier Area (TFA) for each $N$-point time sequence of the historical abnormality-free data set. For each time sequence the TFA is defined as the sum of all Fourier coefficients:

$$TFA(t) = \sum_{n=1}^{N} \Re(n) + \Im(n)$$

(7)

Then, the relative difference of successive TFAs, called the DTFA, is computed for each $t$: 
\[ DTFA(t) = 100 \cdot \left[ \frac{TFA(t) - TFA(t-1)}{TFA(t-1)} \right] \]  

(8)

The DTFA(t) is an indicator of the magnitude of change in traffic conditions from interval t-1 to t. The DTFAs are used to determine the maximum absolute DTFA for abnormality-free conditions, which is used in the next section to detect data abnormalities.

The minimum and maximum real and imaginary Fourier coefficients over all N-point time sequences of the historical data set are used to define the Feasible Fourier Zone (FFZ) for abnormality-free data. The FFZ represents the region in which the coefficients of the Fourier transforms of an abnormality-free time sequence reside with high likelihood. It implies that if the Fourier coefficients lie outside the zone, the associated data is likely to possess an abnormality. Hence, the minimum and maximum coefficient values are used to determine the boundaries of the FFZ. The FFZ is used by the data correction heuristic.

**Detection of Data Abnormalities**

The DTFAs computed for real-time streaming data are used to identify a data abnormality. The Fourier transform for the sequence of N data points at time t is computed. If \(|DTFA(t)| > \lambda\), where \(\lambda\) is a threshold value, a data abnormality is detected. \(\lambda\) is the maximum absolute DTFA observed in the training phase. This approach can adapt to demand changes over a period of time through a retraining step. During the retraining, a new \(\lambda\) reflecting the changed demand conditions is computed using recent data and becomes the basis for the abnormality detection criterion.

**Distinguishing between Data Faults and Possible Incidents**

If a data abnormality is detected, four criteria are used to differentiate between data faults and possible incidents. However, as stated earlier, the confirmation of the presence or absence of an incident requires additional criteria [6]. The first two criteria are based on the relationships between the fundamental traffic flow parameters, speed S, density D, and volume V. While density cannot be measured directly, it can be inferred from occupancy. The remaining two criteria identify data faults in a straightforward manner. These criteria are:

- **Criterion 1:** Speed and density cannot be simultaneously equal to zero. If \(S \to 0\), then \(D \neq 0\) as D tends to the jam density. If \(D \to 0\), then \(S \neq 0\) as it tends to the free flow speed. If this criterion is not satisfied, the abnormality is classified as a data fault. If it is satisfied, a possible incident is indicated.

- **Criterion 2:** The Fault Identification Ratio (FIR) cannot be significantly different from 1:

\[ FIR = \frac{V}{S \cdot D} \]  

(9)

This is due to the traffic flow identity \(V = S \times D\). Hence, if the FIR is significantly different from 1, the abnormality is classified as a data fault. While ideally the traffic flow identity should be conserved, in reality the FIR may differ from 1 due to the mechanism to measure density using occupancy. Density is inferred from the occupancy as follows:

\[ D = \frac{O}{100 \cdot (L_v + L_t)} \]  

(10)
where O is the occupancy in percent, \( L_v \) is the typical vehicle length, and \( L_l \) the loop length. While \( L_l \) is a constant, vehicles have varying lengths requiring \( L_v \) to be an average measure, introducing some approximation in inferring D. Hence, the FIR may not be exactly equal to 1 for a traffic system. Fault-free conditions are likely to exist when the FIR = 1 ± \( \varepsilon \), where \( \varepsilon \) is a threshold value whose magnitude, while small, depends on the particular traffic network being analyzed (based on the vehicle mix). When this criterion is violated, the abnormality is classified as a data fault.

**Criterion 3**: Volume, speed, and occupancy cannot be higher than their maximum possible values in a traffic system. If they are, the abnormality is identified as a data fault.

**Criterion 4**: If volume, occupancy, and speed in successive intervals are equal, the abnormality is identified as a data fault. This criterion implies a jammed detector repeating past data. It can be checked even before computing the DTFAs. Hence, it can be performed in conjunction with the procedure to detect an abnormality.

**Correction of Data Faults**

When an abnormality is identified as a data fault, a data correction heuristic is activated to correct the erroneous data. Its primary objective is to predict the likely correct data. Consequently, it could also be used to predict data values for the near- or medium-term future. Such a capability is significant when detectors malfunction in a traffic system with automated operational control.

The procedure begins with the last data point added to a time sequence being identified as a data fault. This leads to the construction of a N-point time sequence with its last data point deliberately set equal to zero, called a faulty time sequence. This time sequence is updated using a “correct” data point, which is predicted by the heuristic using a fault time sequence which has its first N-1 points equal to zero and the last one predicted by the heuristic. The faulty time sequence is added to the fault time sequence to obtain the corrected sequence. Due to the additive property, the addition of two transforms is equivalent to the transform of the addition of the two sequences. Hence, if the Fourier transforms of the faulty and fault time sequences are added and reversed to a time sequence, the resulting time sequence will have its first N-1 data points equal to the original ones and the last one (Nth) equal to the predicted value.

A key advantage of the proposed approach is that the effect of adding just one data point to generate the corrected time sequence influences the coefficients of all frequencies. Thereby, if the proposed data point is incorrect, its Fourier coefficients lie outside the FFZ. This provides a simple and direct interpretation for the proposed Fourier-based heuristic. The heuristic exploits some convenient characteristics of Fourier transforms. First, we define a *unitary* fault time sequence as a fault time sequence with the last data value equal to one. Its advantage is that the associated Fourier coefficients are obtained in a straightforward manner:

\[
\text{Re}(n) = \frac{1}{N/2} \cdot \cos(2\pi n \cdot v(n))
\]  
\[
\text{Im}(n) = \frac{1}{N/2} \cdot \sin(2\pi n \cdot v(n))
\]
If the predicted last data point has a value $\beta$, the Fourier coefficients of the associated fault time sequence can be obtained using the multiplication property, and (11) and (12):

\[
\text{Re}(n) = \beta \cdot \frac{1}{N/2} \cdot \cos(2\pi t \cdot \nu(n)), \; n = 1, 2, \ldots, N
\]

(13)

\[
\text{Im}(n) = \beta \cdot \frac{1}{N/2} \cdot \sin(2\pi t \cdot \nu(n)), \; n = 1, 2, \ldots, N
\]

(14)

The Fourier coefficients of the corrected time sequence are obtained by adding the computed Fourier transforms of the faulty time sequence to the Fourier transforms of the fault time sequence with value $\beta$:

\[
\text{Re}(n) = \text{Re}_f(n) + \beta \cdot \frac{1}{N/2} \cdot \cos(2\pi t \cdot \nu(n)), \; n = 1, 2, \ldots, N
\]

(15)

\[
\text{Im}(n) = \text{Im}_f(n) + \beta \cdot \frac{1}{N/2} \cdot \sin(2\pi t \cdot \nu(n)), \; n = 1, 2, \ldots, N
\]

(16)

where $\text{Re}_f(n)$ and $\text{Im}_f(n)$ are, respectively, the real and imaginary coefficients corresponding to the faulty time sequence.

We now define a transgression. A transgression refers to any violation of the FFZ. The Total Transgression (TT) at time $t$ is equal to the sum of the absolute transgression values of all the real and imaginary coefficients:

\[
TT(t) = \sum_{n=1}^{N} \{ h_{\text{Re}}(n) + h_{\text{Im}}(n) \}
\]

(17)

where,

\[
h_{\text{Re}}(n) = \begin{cases} 
|\text{Re}(n) - g_{\text{Re}(n)}|, & \text{if transgression of real coefficient at } n \\
0, & \text{if no transgression of real coefficient at } n 
\end{cases}
\]

and,

\[
g_{\text{Re}}(n) = \left\{ \begin{array}{c} \max_{\text{min}} \text{Re}_{\text{FFZ}}(n) 
\end{array} \right.
\]

where $\text{Re}_{\text{FFZ}}(n)$ is the boundary of the real coefficients of the FFZ obtained in the training phase, implying that $g_{\text{Re}}(n)$ is the maximum real coefficient if the maximum value is violated or the minimum real coefficient if the minimum value is violated. $g_{\text{Im}}(n)$ and $h_{\text{Im}}(n)$ can be similarly defined.

Initial experiments showed that in order to maximize the prediction accuracy, only the most recent data points should be used for the determination of the FFZ. This led to the introduction of the recent FFZ, denoted by $\text{FFZ}$. To perform a prediction at time $t$, $\text{FFZ}(L)$ is determined using
time sequences from \((t - L)\) to \((t - 1)\), where \(L\) represents the number of recent past data points. Hence, \(FF\bar{Z}(L)\) is used instead of \(FFZ\) (which considers all time sequences).

The optimal fault time sequence to be added to the faulty time sequence requires the determination of \(\beta\). Two characteristics of non-faulty data are used to determine the optimal value of \(\beta\). A time sequence which does not possess a fault: (i) has minimal transgressions, and (ii) exhibits small DTFAs in small time intervals. These two characteristics provide the basis for the objective function to determine \(\beta\):

\[
\min \left( \frac{TT(t, \beta)}{TT_{\text{max}}} \right)^2 + \left( \frac{DTFA(t, \beta)}{DTFA_{\text{max}}} \right)^2, \quad \beta = 1, 2, \ldots, \beta_{\text{max}}
\]

where \(\beta_{\text{max}}\) is the maximum value that \(\beta\) can take. For example, if the procedure is performed for speed, then \(\beta_{\text{max}}\) will be slightly higher than the free flow speed. \(TT(t, \beta)\) and \(DTFA(t, \beta)\) indicate that \(TT\) and \(DTFA\) are different for different values of \(\beta\). \(TT_{\text{max}}\) and \(DTFA_{\text{max}}\) are the maximum values of \(TT(t, \beta)\) and \(DTFA(t, \beta)\), respectively, for the corrected time sequence with \(\beta = 0, 1, \ldots, \beta_{\text{max}}\). The maximum values are used in the denominator to normalize \(TT(t, \beta)\) and \(DTFA(t, \beta)\), as \(TT\) and \(DTFA\) have very different magnitudes. The first term under the square root corresponds to the first characteristic stated for non-faulty data, and the second corresponds to the \(DTFA\) characteristic. The optimal \(\beta\) obtained from (18) is estimated as the correct data point for that \(N\)-point data sequence.

The procedure to determine the optimal \(\beta\) value using equation (18) is computationally efficient since no reversing of the Fourier transform is required. Hence, the proposed methodology enables the continuous monitoring of the network as part of an automated on-line architecture. It also enables more effective detection of incidents [6].

**EXPERIMENTS USING FIELD DATA**

Experiments conducted to derive insights on the proposed methodologies to detect and correct data faults are discussed in this section. The analysis involves investigating the sensitivity of various parameters used in the methodology and the effectiveness of these methodologies in practice. Field data collected over a 24-day period on a freeway in Hayward, California, and over a six-day period on the Athens, Greece, urban arterial network are used for this purpose. The Hayward field data (www.path.berkeley.edu/FSP), in contrast to the Athens data, provides incident data in addition to the standard traffic flow data. Hence, the experiments to detect data faults and distinguish between data faults and possible incidents are conducted using the Hayward data. The experiments to analyze the data correction heuristic use both data.

**Description of the Data**

The Hayward data includes loop detector volume, speed, and occupancy values aggregated every minute for a 24-day period. It was collected for a 9.2 mile monitored freeway segment, having a range of road cross-sections (3 to 5 lanes, and with or without shoulders). Data was collected for two peak time periods: 5:00-10:00 hrs. and 14:00-20:00 hrs. The Athens data includes traffic volumes (with \(\tau = 90\) seconds). Athens, with a total population of more than 4 million people within a 400 km\(^2\) area, exhibits extremely variable traffic demand leading to complex traffic flow patterns. Therefore, although both data were used for the data correction experiments, emphasis
is given to experiments using the Athens data in order to draw deeper insights on the performance of the proposed heuristic.

**Fault Detection Experiments**

The fault detection methodology is analyzed by deliberately introducing data faults to the volume, occupancy, and speed time sequences every 4 minutes. Introducing faults independently and periodically ensures that a fault could not be correlated or amplified by a previous one. Twenty data fault scenarios were explored as shown in Table 1. Each scenario was tested on a minute-by-minute basis for the 24 days to replicate continuous monitoring. Hence, each fault scenario had 15,240 experiments out of which 3,792 represented faulty conditions. The “randomly in a range” columns in the table imply that data faults are introduced by randomly specifying a value in the proposed ranges (of volume, occupancy, or speed).

The fault scenarios described in Table 1 represent different levels of fault likelihood and severity. The most severe and likely scenarios are 1, 2, 3, 8, 9, 14 and 15. Scenario 1 represents the case of a jammed detector that repeats the same data over and over again. The others (2, 3, 8, 9, 14, and 15) represent malfunctioning detectors from which erroneous data of small magnitude is received. As discussed earlier, a detector malfunction causes certain types of errors, the most likely of which are the aforementioned scenarios.

The other scenarios (4-7, 10-13, and 16-20) are more conservative and primarily explored to investigate the limits of the methodology. Some relate to unrealistically large data error ranges, which are not likely in practice. Examples include scenarios 4, 5, 10, and 16. Others relate primarily to detector hardware imprecisions that generate faulty and/or systematically biased data, requiring hardware fault diagnosis and repair. Hence, they correspond to faulty detectors even before installation at the field sites requiring the associated vendors to address their inaccuracies. Thereby, they are not directly relevant in the context of fault tolerance issues arising for the on-line traffic control architecture where temporary and/or intermittent detector malfunction is the focus. Examples of these scenarios are 11-13 and 17-20. Most detector accuracies range from 85% to 90%. The performance of the fault detection methodology is benchmarked using the same measures used for incident detection: detection rate (DR) and false alarm rate (FAR). It should be noted here that in some cases the methodology recognizes a fault, but is unable to specify whether it lies with volume, occupancy, or speed data. Such faults are labeled as *fault-in-all* time sequences as faults are identified with all of them. Scenario 1 is a good example of this possibility. The implication of this is that the associated data correction heuristic attempts to correct all of them (volume, occupancy, and speed) raising issues of computational efficiency and small losses in accuracy. However, as will be illustrated later, these issues are not practically significant.

Figure 1 illustrates the detection rates for the twenty data fault scenarios, each averaged over the 24 day Hayward data period. The performance of the Fourier-based methodology is very robust for the most likely scenarios, with a DR of 100% and a FAR of 0% in all cases. For the less likely or less relevant faults, the DR averages 76.14% and 87.14% (when fault-in-all are included). A very detailed discussion on the performance of each of these scenarios is provided in [5]. The associated trends are consistent with the expectations from the proposed Fourier-based methodology, and provide deeper insights on it. For example, as expected, the DR for scenario 4 is better than that of scenario 5, though both scenarios are not likely to occur in
practice. The methodology is more sensitive to speed-related faults but less sensitive to occupancy-related faults, and even less to volume-related ones. In terms of differentiating between data faults and possible incidents, there were only 7 cases out of the 304,800 experiments in which a fault was misclassified as a possible incident. However, in none of these cases an incident was confirmed after identifying a possible incident.

**Data Correction Experiments**
The performance of the data correction heuristic is measured using the standard error (SE). The standard error for a predicted data point $m$ is obtained using the correct value for the deliberately faulty data (from the actual data available) and the data points predicted by the heuristic:

$$SE(m) = \frac{|correct(m) - predicted(m)|}{correct(m)}$$  \hspace{1cm} (19)

The average of the SE over $k$ predicted data points, called $ASE(k)$, is used to analyze the performance across $k$ points:

$$ASE(k) = \frac{\sum_{m=1}^{k} SE(m)}{k}$$  \hspace{1cm} (20)

**Preliminary Insights and Sensitivity Analyses**
Preliminary experiments were performed to test the sensitivity of the data correction heuristic to the number of Fourier points $N$. These experiments were performed using one day of historical abnormality-free data to define the FFZ. The performance of the heuristic for $N = 2, 4, \ldots, 512, 1024$ was explored. It was found to perform best for $N = 256$ though the ASE is not sensitive to $N$ for the range $N = 64$ to 256. Smaller $N$ (2 to 32), as well as larger ones (512, 1024), lead to some deterioration for opposite reasons; too few or too many Fourier points. Hence, $N = 64, 128,$ or 256 can be considered as optimal choices, providing tradeoffs in terms of computational efficiency ($N = 64$ is preferred) and performance ($N = 256$ is preferred).

The preliminary analysis of the data correction heuristic used the first day of data for training and the remaining data for the analysis of its performance. While the heuristic performed well in general, there were cases of high ASE, attributable to insufficient training. This implied that the first day data did not adequately represent all possible conditions in the network. To address this issue, a continuous learning feature was introduced by continuously updating the FFZ. This enabled the consideration of all possible conditions up to the time each experiment was performed. The results initially suggested improved performance, which deteriorated with additional training. This is because the continuous training increased the spread of the FFZ, reducing the ability of the heuristic to predict robustly.

As discussed earlier, the preliminary experiments suggest that only the most recent conditions should be used to determine the FFZ, leading to the introduction of $\overline{FFZ}(L)$. The sensitivity of the heuristic to different values of $L$ was explored. It was found to perform best under $L = 5$. Though other values ($L = 2, 3, 40, 50, 100, 150$) were tested, it was observed that in general, the performance improves with decreasing $L$ until about $L = 5$. 
The sensitivity analysis to the number of Fourier points \( N \) was performed again for the modified heuristic. The modified \( \overline{FFZ} \) based data correction heuristic showed little sensitivity to \( N \). The optimal number of Fourier points remained equal to 256. However, the difference in performance is practically insignificant, implying insensitivity to \( N \). This has critical implications for the practical deployment of the methodology, since \( N \) needs to be minimized to more robustly identify faults and incidents. Also, a low \( N \) is computationally efficient vis-à-vis real-time operations. Hence, \( N = 2 \) can be favorable from all relevant perspectives.

Smoothing of raw data is an important aspect of the procedure, as raw data has high levels of volatility. The performance of the heuristic improves as the exponential smoothing parameter value decreases. By decreasing it, more high-frequency components are filtered leading to smoother data. However, tradeoffs arise while smoothing data. Since smoothed data masks information, it reduces the performance of the fault/incident detection methodologies while improving the accuracy of data prediction. However, the data correction heuristic is not sensitive to the smoothing parameter within a certain range. Hence, the parameter value at the higher end of this range can be used to enhance the performance of the fault detection methodology.

**Discussion of Results**

The \( \overline{FFZ}(L) \) approach was used to perform experiments using both Athens and Hayward data. The Hayward data experiments were performed every 60 seconds. They were conducted to predict the data (Volume, Occupancy and Speed) for the next 5 minutes \((k = 5)\), 10 minutes \((k = 10)\), and 20 minutes \((k = 20)\). As expected, while \( k = 5 \) leads to the best performance, the other experiments were performed to provide insights on the heuristic performance for larger periods of prediction \((10 \text{ and } 20 \text{ minutes})\). As illustrated in Table 2, the heuristic achieved the best performance for speed data errors, and the worst for occupancy related ones. The occupancies on this freeway are very low \((1\% \text{ is a common situation})\) for many durations during any 24-hour period. Hence, a small difference in the prediction may manifest as a very large SE. Therefore, practically, the performance vis-à-vis occupancy is robust for the typical periods of interest.

The Athens data, in contrast to the Hayward data, is continuous and includes both peak and off-peak periods, enabling us to derive insights on the performance of the heuristic under more general traffic conditions. Using the Athens data, experiments were performed every 90 seconds on all six days for three different network links \((1, 5, \text{ and } 65)\) that are representative of the different traffic conditions. Link 5 has significantly lower volumes than 65, and significantly different traffic patterns. It has one widespread peak each day while link 65 has two. Link 1, the third test link, has characteristics similar to link 5 with higher average daily volume \((\approx 22,000 \text{ vehicles/day instead of } 18,400 \text{ vehicles/day})\), and more fluctuations. Link 65 has an average daily volume of approximately 30,000 vehicles/day.

The experiments were conducted to predict data for the next 7.5 \((k = 5)\), 15 \((k = 10)\), and 30 \((k = 20)\) minutes. Figure 2(a) depicts the heuristic performance for 3-hour aggregated data for link 65. The ASE is aggregated every 3 hours and 3-hour averages are presented. The heuristic is found to perform well except during night hours (from midnight till 6:00 AM) for all days as indicated by the peaks. This is because the volumes during these periods are very low, making traffic conditions almost random. Figure 2(b) depicts the heuristic performance for link 65 by
excluding night hours, leading to significantly lower average standard errors. Figures 3(a) and 3(b) illustrate the performance for links 1 and 5, respectively, after excluding the night hours.

The heuristic demonstrates an average ASE equal to 6.92%, 9.41%, and 11.98% for 7.5-, 15-, and 30-minute predictions, respectively. Existing forecasting techniques in this arena usually provide 15-minute single interval predictions. Since there is no base for comparison at the minute-to-minute prediction level to obtain a general feel for the performance of the proposed methodology, the same experiments were performed using a simple average projection technique as a benchmark. According to this technique, the average of the most recent values was projected to the future. This technique yielded ASEs equal to 28.78%, 29.31%, and 30.47%, respectively, for $k$ equal to 5, 10, and 20. The robustness of the data correction heuristic is highlighted by its greater accuracy compared to the benchmark.

CONCLUDING COMMENTS

This paper proposes a Fourier-based fault tolerant framework for detecting and correcting data errors due to malfunctioning detectors. It is intended for use in an automated on-line architecture for real-time route guidance in traffic systems equipped with advanced sensor and information dissemination devices. It can also be used to detect incidents. The framework analyzes data directly and avoids elaborate modeling. The proposed methodologies do not require elaborate training, and use little data for training. Hence, the training is straightforward, with the Fourier-related parameters being simple to compute. However, these parameters need to be re-computed if major demand/supply changes occur over short time periods. Also, the parameters are not transferable to new sites that have significantly different system characteristics, but can be computed easily if the relevant data is available. Another issue vis-à-vis data is the requirement of abnormality-free data for training purposes. The data correction heuristic is a critical component of the automated on-line control architecture. The proposed Fourier-based methodology is computationally very efficient. It is robust for most practical scenarios, even under the relatively more volatile Athens traffic data as well as under a wide range of traffic conditions. However, the heuristic is not robust for low volume/congestion traffic situations, typically encountered during night. This does not affect its value for on-line traffic control architectures which are likely to be employed primarily for congested traffic conditions. Also, the heuristic performs better for near-term (5 to 7.5 minutes) or medium-term (10 to 15 minutes) predictions than for long-term ones (20 to 30 minutes). This is a general limitation of any prediction methodology.

A key contribution of the research is that the same (Fourier-based) methodology is used to detect faults and incidents, correct data faults, and predict traffic flow parameters for the near-term future on a minute-to-minute basis. This has significant implications for an automated on-line control architecture for real-time traffic operations. The use of a unified (Fourier-based) approach to perform several critical tasks enhances computational efficiency vis-à-vis real-time operation and simplifies the architecture. A key aspect of the methodology is the ability to predict traffic conditions for the near-term future on a minute-to-minute basis efficiently. This synergistically aids the deployment of the on-line architecture as most existing efficient traffic prediction approaches can provide predictions over larger time intervals only (typically 15 minutes).
ACKNOWLEDGEMENTS

This material is based upon work supported by the National Science Foundation under Grant No. CMS-9702612. Additional funding is provided through the Joint Transportation Research Program at Purdue University. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation, the Federal Highway Administration and the Indiana Department of Transportation. The authors would like to thank Anthony Stathopoulos for providing the Athens traffic data and George Mylonakis for providing software to compute the Fast Fourier Transforms.

REFERENCES


LIST OF TABLES AND FIGURES

TABLE 1 The Scenarios for the Fault Detection Experiments

TABLE 2 Results of the Data Correction Experiments for Hayward Data

FIGURE 1 Detection rates for the fault detection experiments

FIGURE 2 Results of the data correction experiments for Athens data; Link 65

FIGURE 3 Results of the data correction experiments for Athens data; Links 1 and 5
<table>
<thead>
<tr>
<th>Scenario Number</th>
<th>Volume (vph)</th>
<th>Occupancy (%)</th>
<th>Speed (mph)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Specific value</td>
<td>Randomly in a Range</td>
<td>Under-estimated</td>
</tr>
<tr>
<td>1</td>
<td>Previous Interval Volume</td>
<td>Previous Interval Occupancy</td>
<td>Previous Interval Speed</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>0 – 10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>0 – 500</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>0 – 800</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>50 %</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>50 %</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>0 – 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>15 – 30</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td></td>
<td>30 %</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td></td>
<td>40 %</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td></td>
<td>40 %</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td></td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td></td>
<td>0 – 5</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td></td>
<td>0 – 30</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td></td>
<td>25 %</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td></td>
<td>30 %</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td></td>
<td>30 %</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td></td>
<td>60 %</td>
<td></td>
</tr>
</tbody>
</table>

**TABLE 1** The Scenarios for the Fault Detection Experiments
<table>
<thead>
<tr>
<th>Prediction Range</th>
<th>ASE for Volume (%)</th>
<th>ASE for Occupancy (%)</th>
<th>ASE for Speed (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$k = 5$</td>
<td>13.2</td>
<td>15.1</td>
<td>4.6</td>
</tr>
<tr>
<td>$k = 10$</td>
<td>14.0</td>
<td>15.9</td>
<td>5.1</td>
</tr>
<tr>
<td>$k = 20$</td>
<td>15.2</td>
<td>17.4</td>
<td>6.0</td>
</tr>
</tbody>
</table>

**TABLE 2** Results of the Data Correction Experiments for Hayward Data
FIGURE 1 Detection rates for the fault detection experiments.
FIGURE 2  Results of the data correction experiments for Athens data; Link 65.
FIGURE 3  Results of the data correction experiments for Athens data; Links 1 and 5.