Cross-Correlation Based Vehicle Feature Extraction by Magnetic Wireless Sensor Networks

Kejiang Xiao[1, 2], Rui Wang[1], Lei Zhang[3], and Lei Xu[1]
[1] School of Computer and Communication Engineering, University of Science and Technology Beijing, Beijing 100083, China
[2]State Grid Information & Communication Company of Hunan Electric Power Company, Changsha 410007, China
[3] Department of Computer Science, Georgia State University, Atlanta 30303, USA
Email: xkj2006xkj@126.com; wangrui@ustb.edu.cn; lzhang14@student.gsu.edu; xulei_717@163.com

Abstract —Due to the limitation of sensor size and cost, in order to implement vehicle detection by magnetic wireless sensor network, it attaches importance to the two important problems, vehicle detection and feature extraction, with facing the challenge of low signal-to-noise ratio and limited resources. In this paper, we propose a mechanism to accomplish vehicle monitoring, namely CBNP. In this mechanism, we propose a cross-correlation based vehicle detection algorithm to accurately detect vehicle presence. To accurately extract response feature, we propose and prove a theorem which depicts the relation between feature point and its cross-correlation result. Based on the theorem, we proposed a Cross-Correlation Based Feature Extraction (CBFE) algorithm to accurately locate vehicle feature points under relative strong noise. Simulation indicates that CBNP outperforms traditional methods on both vehicle detection rate and feature extraction accuracy. In meaningful conditions, the detection accuracy is over 90% and the deviation rate of feature extraction is within 10%.

Index Terms—Wireless sensor networks, feature extraction, vehicle monitoring

I. Introduction

Vehicle speed estimation is important for Intelligent Transportation System. Individual vehicle speed is the distance travelled divided by the time taken to traverse that distance. At present there are several ways to measure individual vehicle speed, such as inductive loop and radar device. Due to its non-intrusive nature, Wireless Sensor Networks [1] (WSN) have recently been proven to be ideal solutions for vehicle speed estimation. As vehicle monitoring systems [2], [3] can improve urban transportation efficiency, much attention has been paid to this field. In WSN regime, there are several methods to measure individual vehicle speed. One kind of methods is to measure vehicle speed with the insistence of on-vehicle device, such as Global Position System (GPS) device and Radio Frequency Identification (RFID) tag [4]. Another kind of methods is to measure vehicle speed with only the roadside sensors. In this paper we focus on the later kind of method.

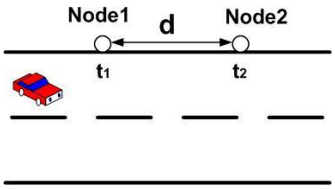
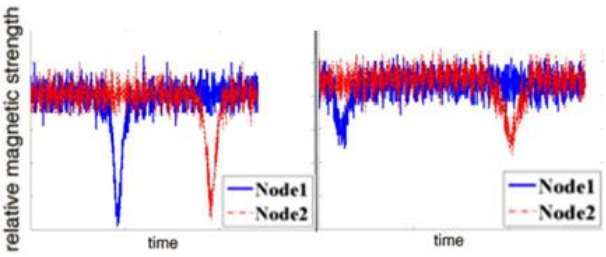


Fig. 1. Typical WSN speed measurement

Fig. 1 illustrates a typical speed measurement scene in WSN. Vehicles speed can be calculated by $v = d/|t_1 - t_2|$. But there are several prerequisites. First, as t_1 and t_2 are the report vehicle appearance time instant of a vehicle, they are the local time of these two nodes. So the local timer of Node1 and Node2 should be synchronized at a relatively high level (microsecond level). Second, Node1 and Node2 should accurately report the time instant when a vehicle "exactly" passes that node, which is a challenge because the background noise introduces uncertainty and response signal in each node is difficult to be identical.



(a) (b)

Fig. 2. Signals in vehicle speed measurement.

Fig. 2 illustrates signals sampled when a vehicle runs in different lanes in Fig. 1. It is obvious that the useful signal in Fig. 2 (a) is relatively stronger than that of Fig. 2 (b). It is relatively easy to identify t_1 and t_2 in sample Fig. 2(a), but it is difficult to identify t_1 and $t_2$in sample Fig. 2(b).

To realize a reliable and accurate WSN based vehicle speed measurement system, there are two challenges. The first challenge is to identify each vehicle from the sampled signal, because if a vehicle is not recognized, the speed of that vehicle will not be calculated at all. The second challenge is to accurately calculate the time

Manuscript received October 12, 2015; revised April 8, 2016.
This work was supported in part by the National Natural Science Foundation of China under Grant No. 61379134.
Corresponding author email: wangrui@ustb.edu.cn.
doi:10.12720/jcm.11.4.349-357

instant when a vehicle exactly passes a sensor node, even when the noise is relatively strong.

In this paper we propose and verify a vehicle speed measurement solution based on WSN, named CBNP. We consider our contribution as follows. First, we proposed and realized a vehicle detection algorithm. The detection accuracy is over 90%. Second, we propose and verify a vehicle response waveform feature extraction algorithm, named Cross-Correlation Based Feature Extraction (CBFE), which can accurately report the vehicle appearance time instant. Third, on-road experiment verified the proposed solution, which is applicable in real environment.

The rest part of this paper is organized as follows: Section 2 discusses the related work. Section 3 and 4 propose the vehicle detection algorithm and feature extraction algorithm respectively. Section 5 provides verifications and experiments. Section 6 concludes the paper.

II. Related Work

Magnetic WSN based vehicle monitoring applications have been widely explored and many solutions have been proposed [5]-[8]. Although many innovations have been proposed to accomplish vehicle monitoring task, there are still challenges in magnetic vehicle detection which are not well resolved. The detection quality is still severely influenced by ubiquitous background, especially when the Signal-Noise-Ratio (SNR) is relatively low. Also due to the noise or low SNR, the vehicle response feature cannot be accurately extracted.

Vehicle detection by magnetic sensor network was proposed in [6]. In this paper, a simple threshold based vehicle detection algorithm was proposed. It states vehicle will be announced when sensor reading exceeds a certain threshold. In [7], the improved part is the simple threshold in vehicle detection can be dynamically adjusted by a Constant False-Alarm Rate (CFAR) detector. Based on the detection results, the sensor node extracts three vehicle response features: time duration that the sensory data exceeds the threshold, maximum sensor reading and its time instant. Researchers in [9] proposed similar Closest Point of Approach (CPA) algorithm to accomplish detection task and extract vehicle feature points from the raw signal with the help of an adaptable threshold. However, these works didn't pay enough attention to noise, so the slight decline of SNR may lead to poor system performance.

In [8], the sampled sensory data is first processed by a moving average statistics (MAT) to reduce noise and then use a threshold to determine whether a vehicle appears. After that, vehicle feature can be calculated. However, this method is also vulnerable to low SNR.

In [5], the noise is reduced by a digital filter (low pass filter, LPFT), and the detection task is then performed. In [10], the raw signal is first input into a limiter and then a low pass filter to reduce noise, after that, decimator, moving statistics, CFAR and energy estimator will be imposed on the signal and finally some features of vehicle response signal, such as duration, energy, will be calculated. However, the first two steps, limiter and low pass filter, cannot effectively reduce noise in all cases. Furthermore, the low pass filter will result in phase deviation, so the following steps may build on an inaccurate input.

All the aforementioned methods are mainly based on three basic mechanisms: 1. Direct thresholding: A threshold is used to detect vehicle. 2. Moving Average and Thresholding (MAT): MAT uses an average of neighbouring data to replace original data so that noise disturbance can be reduced. Then a threshold is used to detect vehicle. 3. Low Pass Filter Thresholding (LPFT): LPFT uses a low pass filter to process raw data and then a threshold is used to accomplish detection tasks. Direct thresholding has been proven to be unfit for real application environment. To some extent, MAT and LPFT can be implemented in real application. On the basis of vehicle detection algorithm, feature extraction algorithm locates feature point. The feature extraction accuracy directly decides high-level information fusion quality; however, noise always results in inaccurate feature calculation.

As for Feature Extraction, an improved cross-correlation algorithm [11] is presented for high-precision pipeline leak detection based on wavelet transform and energy feature extraction. But it does not provide ideal response model and apply for vehicle detection. [12] develops a novel framework for feature extraction based on a combination of Linear Discriminant Analysis and cross-correlation. The frequency domain signal is then cross correlated with predefined classes of ECG (Multiple Electrocardiogram) signals, in a manner similar to pattern recognition. But it is not suitable the scenario of this paper.

III. Cross-Correlation Based Vehicle Detection

In CBNP mechanism, effective feature extraction should be built on accurate vehicle detection, because if vehicle is not detected, feature extraction cannot be triggered at all. In this section we will introduce the cross-correlation based vehicle detection algorithm, then we build vehicle response model, which is critical in cross-correlation calculation.

A. Detecting Vehicles by Cross-Correlation

The existing vehicle detection methods mainly confirm vehicle presences by a threshold. Normally, the raw data will be pre-processed by customized MAT and LPFT algorithms to reduce noise. Then a threshold will be used to decide the vehicle appearance. We argue that in this common framework, vehicle detection cannot be effectively accomplished. Fig. 3 shows a strong response and a weak response sampled in the direction of X-axis when a vehicle passes a sensor node. Ideally, vehicle

detection task can be accomplished by processing the raw magnetic data (Fig. 3a). However, relative strong noise or low SNR (Fig. 3b) always results in high false detection rate and low detection rate.

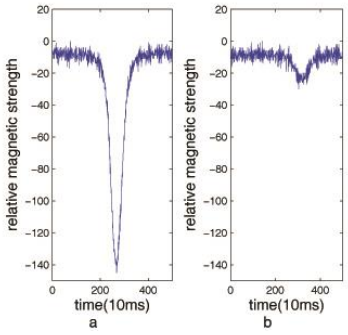


Fig. 3. Different responses caused by vehicles (a.strong b.weak).

We solve the vehicle detection problem from a different perspective. We use a designed filter to reduce SNR of sampled data and detect vehicle according to the filter output. Suppose the system function of our designed filter is $H(jw)$ which is the input and the output is $O(t)=s_0(t)+n_0(t)$. Then the SNR at time instant t can be expressed by $p(t)=s_0^2(t_m)/n_0^2(t_m)$.

As $s_0^2(t)=\frac{1}{2\pi}\int_{-\infty}^{\infty}H(jw)s(jw)e^{jwt}dw$, the noise $n(t)$ is commonly supposed to be Gaussian white noise with a power spectrum of $|H(jw)|^2 \bullet N$ (N is the constant of power spectrum), so

$$p(t)=\frac{\left|\int_{-\infty}^{\infty}H(jw)S(jw)e^{jwt}dw\right|^2}{2\pi N\int_{-\infty}^{\infty}|H(jw)|^2\,dw} \tag{1}$$

According to the Cauchy-Schwarz inequality, the possible $p_{\max}$ happened when $H(jw)=k\left[S(jw)e^{jwt}\right]^*$ (k is any constant). This happens when the correlation between input signal $s_0(t)$ and the system function $h(t)$ is maximized.

We suppose that there is a vehicle response model which is similar to most vehicle responses. We transform the vehicle detection task to the similarity calculation between real vehicle response and vehicle response model. Here we choose cross-correlation [13] as the similarity measurement. Suppose that the vehicle response model is $h(t)$. The real vehicle response is $H(t)$ and $H(t)=s(t)+n(t)$, where $s(t)$ is the ideal vehicle response and $n(t)$ is noise. Here we choose cross-correlation operation (R^*).

As for the vehicle response model, the ideal vehicle response is highly correlated but the noise is irrelevant, thus after cross-correlation calculation, $corr(h(t),s(t))$ will reach peak while $corr(h(t),s(t))$ will remain low. $corr(h(t),H(t))=corr(h(t),s(t))+corr(h(t),n(t))$.

B. Ideal Vehicle Response

In (1), the first and critical step of cross-correlation calculation is the choice of vehicle response model. In this part we will discuss the vehicle response model and then build a vehicle response model.

When a large ferrous object passes an observation point from nearby, the earth magnetic intensity will be changed. Fig. 4 shows a scene in which a vehicle passes an observation point (sensor). We suppose the time instant starts at $T=0$. Then the speed of vehicle running at the horizontal direction is *v*, which recognized as invariable when vehicle is in the sensor's probing range. In addition, the vertical distance between vehicle and sensor signed as *d* and the horizontal distance as *l* (We suppose *l* is large enough). So the vehicle-caused magnetic intensity [14] along the horizontal direction around sensor can be expressed by the following equation:

$$h=m/ur^3 \tag{2}$$

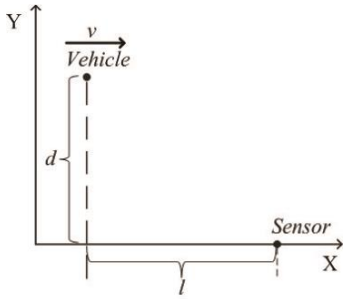


Fig. 4. Vehicle and sensor

In (2), *m* is earth pole strength, which is a constant, *r* stands for the distance between vehicle and sensor, μ states a media constant. At time instant $T=t$, distance between vehicle and sensor can be expressed by:

$$r(t)=\sqrt{d^2+(l-vt)^2} \tag{3}$$

According to (2) and (3), the magnetic intensity at time instant t can be expressed by the following equation:

$$h(t)=m/(u\bullet(d^2+(l-vt)^2)^{\frac{3}{2}}) \tag{4}$$

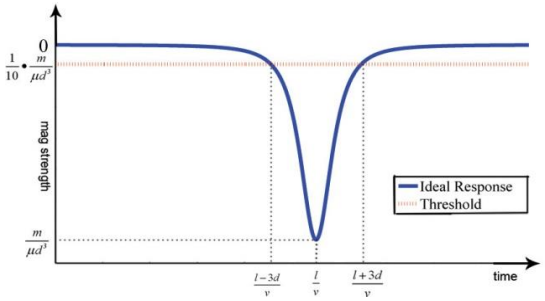


Fig. 5. Ideal vehicle response

So the ideal vehicle response is showed in Fig. 5. From Fig. 5 we can see that the vehicle response is symmetric about $t=l/v$. As the vehicle response is very weak in most time duration, here we use a threshold $\left|t-\frac{l}{v}\right|<\frac{3d}{v}$

(as displayed in Fig. 5) to build vehicle response model. So the vehicle response model can be expressed by following equation (l is removed):

$$h(t)=\begin{cases} m/u(d^2+v^2t^2)^{\frac{3}{2}} & |t|<\dfrac{3d}{v} \\ 0 & other \end{cases} \quad (5)$$

C. Real Vehicle Response

When a real environment is considered, the noise cannot be ignored. Here we suppose noise w to be the Gaussian white noise with zero mean and variance δ^2. So the modelling of noise is to identify the parameter δ^2. We will determine the value of δ in simulation section. So real vehicle response can be expressed by:

$$H(t)=\frac{m}{u\bullet(d^2+(l-vt)^2)^{\frac{3}{2}}}+w(t) \quad (6)$$

D. Threshold-Insensitive Vehicle Detection

According to (1) and (5) we propose cross-correlation based vehicle detection algorithm. The basic idea of algorithm contains four steps described as follows.

Step 1: Compute the real vehicle response $H(t)$.

Step 2: Compute the ideal vehicle response $h(t)$.

Step 3: Compute the cross-correlation $corr(h(t), H(t))$ between the real vehicle response $H(t)$ and the ideal vehicle response $h(t)$.

Step 4: Make comparison between the value of $corr(h(t), H(t))$ and a proper threshold in similarity value to decide whether the vehicle is absence.

Fig. 6 shows the basic idea of the algorithm. It can be shown that the similarity values of those two responses are very close to each other although their raw signal waveforms are very different from each other. With a proper threshold in similarity value, the two vehicles can be detected. Although noise is ignored in this figure, we will demonstrate in simulation section that our algorithm keeps a high detection rate even noise is relative strong.

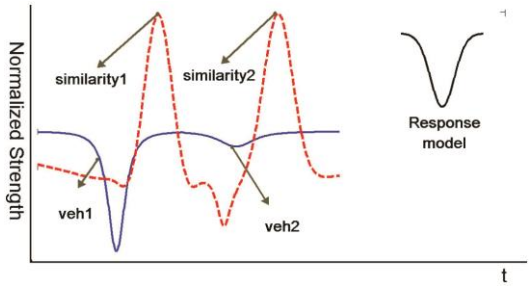


Fig. 6. Similarity values of two different vehicle responses.

IV. Cross-Correlation Based Feature Extraction Algorithm

Vehicle features of a vehicle response can be described by several key points. Among all these feature points, peak point is the most important one, because the peak point can be a beacon for the identification of rest points, furthermore, the peak point indicates the vehicle presence time and response strength. However, the identification of peak point is always disturbed by noise or disturbance. In Fig. 3b, as the noise is relative strong, signal has been severely distorted, so it is difficult to identify the peak point.

In this section we propose a CBFE algorithm to accomplish feature extraction task in CBNP. CBFE can accurately identify peak point even when SNR is relatively low. As in a sensor node, the cross-correlation calculation should be implemented in a discrete form, which is expressed by (7). In (7), f is a sampled data series, g is data series of vehicle response model, g is the time duration of M.

$$f(t)*g=\frac{\sum_{m=0}^{M-1} f(t+m-M)g(m)}{\sqrt{\sum_{m=0}^{M-1} f^2(t+m-M)}\sqrt{\sum_{m=0}^{M-1} g^2(m)}} \quad (7)$$

Based on the cross-correlation properties, we have the following theorem:

Theorem 1: Suppose Given the response model, t_p is the time instant of vehicle response peak, t_s is the time instant of its cross-correlation peak, then, t_s-t_p *is* a constant, and $t_s-t_p=M/2$, where M is the length of response model.

Proof: In appendix.

Fig. 7 shows the relation depicts by theorem1.

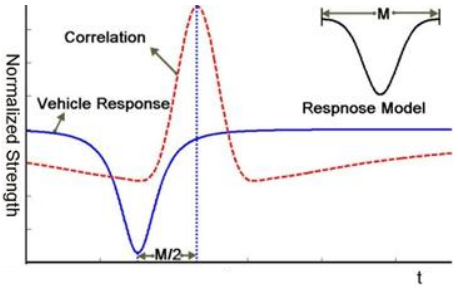


Fig. 7. Time difference between response peak and cross-correlation peak.

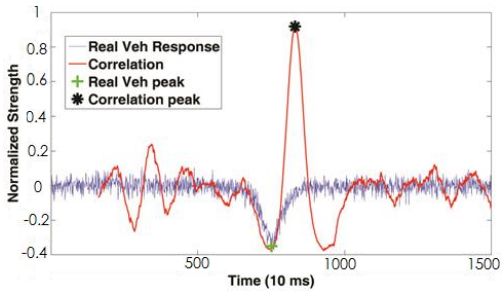


Fig. 8. Real vehicle response and its cross-correlation peak.

Now we propose CBFE feature extraction algorithm. When the cross-correlation peak is located, the corresponding raw data peak can easily be located. This is the basic idea of CBFE algorithm. In the sampled real data, noise is also included and identification of the

lowest point (peak point) is difficult. However, the cross-correlation result is rarely disturbed by noise. Fig. 8 is a cross-correlation result of real vehicle response with noise included. From this figure we may find that, due to relative strong noise, through traditional methods, it is difficult to locate the peak of vehicle response. The cross-correlation output is much smoother, so the cross-correlation peak can be easily located and the vehicle peak can be accurately located.

V. Simulation and Verification

In this section we verify the CBNP mechanism. First we verify the detection accuracy of CBNP by both simulation and experiments. Then we estimate the feature extraction accuracy of CBNP. Finally we check the relation between algorithm performance and SNR to show that our proposed solution deals with lower SNR. In all our simulation and experiments, we choose MAT algorithm and LPFT algorithm as baseline algorithms.

According to (6), the vehicle response is mainly influenced by v, d and ω. In this section we also discuss the relation between performance and v, d, ω. As v and d should be meaningful for monitoring applications, we will numerate all the meaningful values to evaluate performance. We choose two algorithms to compare with, which are MAT and LPFT algorithm.

According to the monitoring application context, several factors should be taken into consideration. As HMC100x [15] is often chosen as magnetic sensor in monitoring applications [3], [5] and their probing range is within 6 meters and the vehicle speed in urban areas is mainly between 20kph to 50kph (about 6m/s to 14m/s), here we use different d (range from 1 to 6, 0.5 as step) and different v (range from 6 to 14, 1 as step) to generate real vehicle responses described in (6). According to the on-road sensory data, we set δ^2 of noise to be 11.5 (unit is HMC1000's resolution). For each run, we specify a value for d and v, then 1000 vehicle responses will be generated and each of them will be processed by three mechanisms (CBNP, MAT and LPFT). Then we give statistical results.

A. Simulation Settings

The MAT algorithm we adopted in our simulation will use the mean of 5 time instants (current, previous two and post two instants) value to replace the value of current time instant, and the mean-replacement operation will be conducted three times. By this "smooth" operation, the uncertainty caused by noise can be reduced. The LPFT filter we adopted is a low pass filter, whose pass frequency is 4Hz and stop frequency is 9Hz. When the cross-correlation based vehicle detection algorithm in CBNP is referred, as we have mentioned, response model should be similar to most vehicle response, so here we choose $d = 4$ and $v = 4$.

B. CBNP Verification

In the verification of CBNP, we first compare our solution with two baseline algorithms by computer simulation. Then we compare detection accuracy by processing real on-road sampled data, the comparative results of two baseline algorithms are also demonstrated.

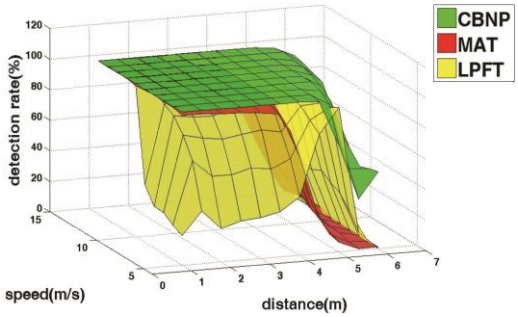


Fig. 9. Detection accuracy vs. speed and distance.

- **Computer simulation**:

To verify the vehicle detection accuracy, we generate vehicle responses with different d (1 to 6, 0.5 as step) and different v (6 to 14, 1 as step). For each d and v, 1000 vehicle responses will be generated and we will calculate the detection accuracy by statistics. When a vehicle response is processed by an algorithm, detection failure will be announced if no vehicle is detected or more than one vehicle is announced. Fig. 9 shows the overall results of simulation. From this figure we can see that CBNP detection accuracy mainly outperforms the rest two algorithms.

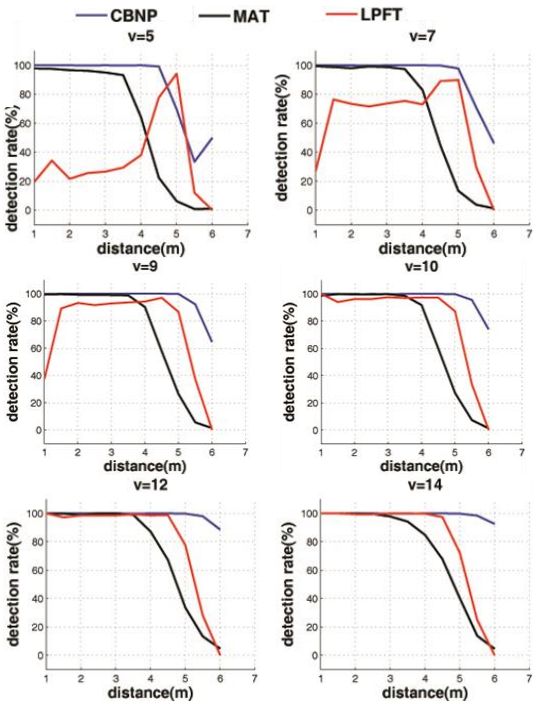


Fig. 10. Detection rate and distance under different speed.

Fig. 10 shows several slices in Fig. 9 with different v. As can be seen from this figure, the detection accuracy will decrease with the increment of distance. However, in on-road monitoring applications, most d of vehicle is within 4m, so the change of d does not severely influence CBNP performance, the detection accuracy is more than 90% when $d < 4$.

- **Processing of on-road sampled data**:

In this part, we use collected data to verify the performance of our proposed algorithm. The on-road settings are demonstrated in Fig. 11. The collection was conducted on a two-lane road, the sensor node was put on the roadside, collected data was forwarded to the sink. The sample rate was set to be 100Hz. A digital video was put aside to catch visual data simultaneously. The ground truth was built on watching the synchronized video. According to our video based statistical result, there was 167 vehicles appeared in our experiment.

Fig. 11. On-road experiments and settings

Fig. 12 is a piece of our collected data, in this figure there are three typical parts. The b and c parts are responses of two target objects, we may see that target response in b is relatively weak. The reason for this phenomenon may because the b target ran on the distant lane while the c target went on the close lane. There was no target appears in a part, so a part consisted of pure background noise. As the feature of noise is relatively stable, we estimate noise parameter based on data in a part. In our verification, we adopt the result that the variance of noise is $\delta^2 = 11.5$.

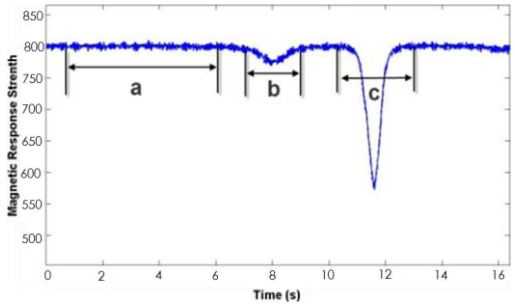


Fig. 12. Raw magnetic data of on-road sensor

After the data was collected, three nodes in which three algorithms are implemented to process the collected data. By changing the thresholds, algorithm performances are recorded, including missing detection rate and false detection rate. The result is demonstrated in Fig. 13. From Fig. 13(a) we can see that, to all the three algorithms, the false detection rate will decrease as the threshold increases. The reason is that, high detection threshold will reduce the uncertainty introduced by background noise but neglect some weak responses such as the target b in Fig. 12. In any reasonable threshold, our proposed CBNP algorithm outperforms the other two algorithms. According to Fig. 13(b), as to the CBNP and LPFT algorithms, the missing detection rate increases as the threshold value are increased. As to the MAT algorithm, with the increment of threshold, algorithm performance changes like a "V" curve. Though in some conditions, our proposed CBNP algorithm has a worse performance than the comparative algorithms, the CBNP has a better overall performance (missing detection and false detection), as described in Fig. 14. Fig. 14 indicates the error detection rate of three algorithms. In this figure, we may conclude that only our proposed CBNP algorithm has a performance of error rate lower than 10%.

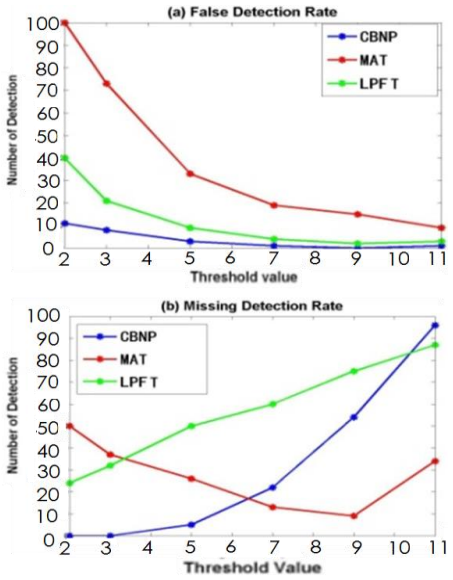


Fig. 13. Detection performance comparison

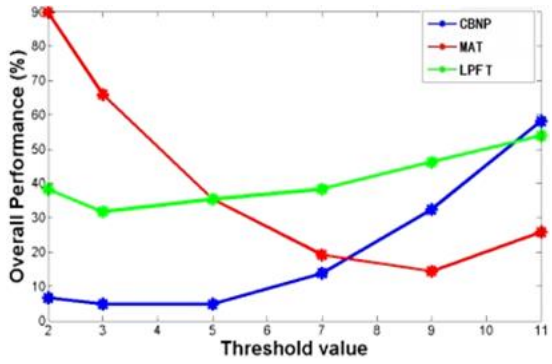


Fig. 14. Overall performance comparison

In this part of verification, we can conclude that CBNP keeps a relatively stable but high detection quality, so the threshold selection is no longer a technical problem, which means CBNP is a threshold-insensitive algorithm.

- **Choice of referential waveform**:

In previous section we mentioned that the choice of response model is a critical step. In this part we demonstrate the criteria of response model choice. As the response model should be similar to most target response, we select a common target response, as displayed in Fig. 15(a). This is a response of a target passing sensor node at a speed of 30km/h and from a distance about 3m. So

this response is a typical target response in traffic monitoring. In Fig. 15(b), we choose a proper response model indicated by dotted line, which is similar to the target response itself. The model itself can be depicted by

$$R(t) = -4287.5 * (12.25 + 0.04t^2)^{-\frac{2}{3}} (-80 < t < 80) \quad (8)$$

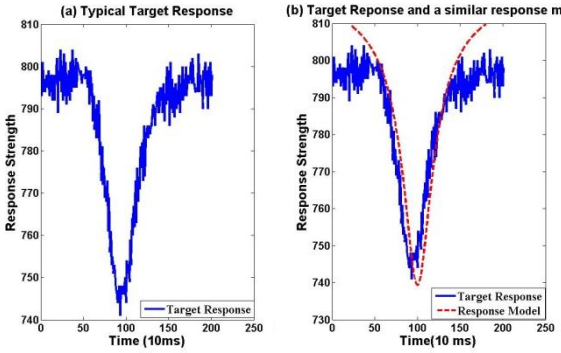


Fig. 15. Building model for a typical response model

In this part, we use $R(t)$ as a baseline, we conducted several detection experiments on the on-road sampled data in the previous section. We demonstrate the detection performance by changing the CBNP response model from $R(t)$, $R(2t)$, $R(3t)$, $2R(t)$, $0.5R(t)$. Table I listed the false detection rate under different response model. From this table we may see that, with different response models, the algorithm has nearly same low false alarm rate. Table II indicates the missing detection rate, with different response models, the missing detection rate increases as the threshold grows up. Table III indicates the overall error rate, by this table we may see that, although response model may vary, the algorithm performance can be guaranteed by a choosing a proper threshold.

TABLE I: FALSE ALARM RATE (FDR)

FDR	2	3	5	7	9	11
R(t)	5.99	2.99	0.60	0	0	0
R(2t)	7.78	1.20	0	0	0	0
R(3t)	7.78	1.20	0	0	0	0
0.5R(t)	5.39	2.40	0	0	0	0
2R(t)	5.99	2.40	0	0	0	0

TABLE II: MISSING DETECTION RATE (MDR)

MDR	2	3	5	7	9	11
R(t)	1.20	1.80	13.77	41.92	74.25	98.2
R(2t)	0	2.99	28.14	68.86	97.01	100
R(3t)	0	3.59	41.91	87.43	100	100
0.5R(t)	0	1.80	13.17	44.91	74.85	100
2R(t)	0	1.80	14.37	46.11	76.65	100

TABLE III: OVERALL ERROR RATE (OER)

OER	2	3	5	7	9	11
R(t)	7.19	4.79	14.33	41.92	74.25	98.2
R(2t)	7.78	4.19	28.14	68.86	97.01	100
R(3t)	7.78	4.79	41.91	87.43	100	100
0.5R(t)	5.39	4.20	13.17	44.91	74.85	100
2R(t)	5.99	4.20	14.37	46.11	76.65	100

In this part, we demonstrated that the response model selection is not strictly related to algorithm performance, but the response model should be within a meaningful range: it should be similar to real target response.

C. Feature Extraction Accuracy

In this simulation we verify the accuracy of peak point feature extraction. For each vehicle response, three versions of vehicle response feature will be calculated individually by LPFT, MAT, and CBNP. All the three results will be compared with the actual peak point. Standard deviation, will be calculated, and the accuracy is measured by deviation rate, where width is the actual response width, which can be calculated according to (5).

Fig. 16 shows the extraction accuracy statistics in our simulation. From this figure we can see that extraction quality of CBFE is better than that of rest two algorithms in almost all situations. According to Fig. 16 we may see that when the d is small, MAT accuracy is similar to CBNP accuracy, but while the d increases, the MAT accuracy will quickly decreases while the CBNP accuracy remains in a relative high quality. LPFT is not so accurate because that the low pass filter will result in a phase difference which means the feature information is distorted. It should be noted that according to Theorem1, the lag between cross-correlation peak and real vehicle peak is exactly $M/2$, which means the deviation rate is 0. While in our simulation the deviation rate is above 0, the reason is that the noise will also account in the cross-correlation result. Luckily the similarity between noise and response model is much smaller than that between model and vehicle response. As in realistic condition, the distance between vehicle and sensor node can be controlled in 4m, so in this case the deviation rate of CBNP is less than 10%.

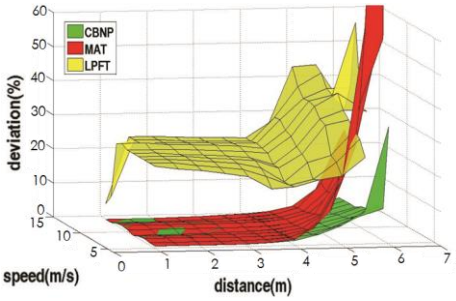


Fig. 16. Feature extraction accuracy

D. SNR Impact on Performance

As we have mentioned before, ubiquitous noise leads to the uncertainty in vehicle feature extraction and low SNR always leads to detection failure. In this part we discuss the relation between SNR and performance.

As we have mentioned before, most existing vehicle detection algorithms do not work well under low SNR. So we will quantify the SNR to measure detection quality and extraction accuracy. It should be noted that in our following discussion, the SNR is calculated according to the simulation result.

Fig. 17 shows the relation between SNR and detection quality (detection rate). Fig. 18 shows the relation

between SNR and extraction quality (deviation rate). As we can see from Fig. 17, detection quality can be assured when the SNR is high enough, as in that case, the uncertainty caused by noise will not lead to detection failure at all. When the SNR decreases detection quality of three methods decrease to nearly zero, but CBNP can work under lower SNR effectively. Even under the same low SNR, CBNP has a better detection quality. Fig. 18 indicates that when SNR is large enough, the feature extraction quality of MAT and CBNP is nearly the same. However, as the SNR decreases, the MAT quality quickly gets worse but CBNP remains a relative high quality. LPFT quality is not high at all because that the low pass filter will introduce a phase difference of feature point.

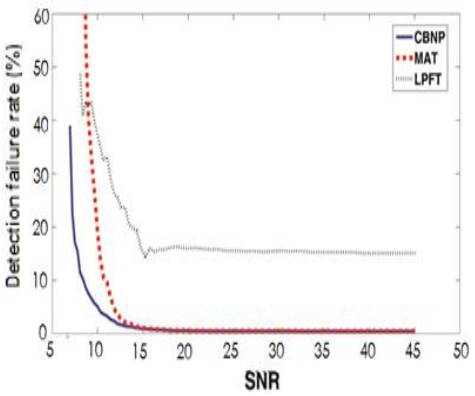


Fig. 17. Detection rate and SNR

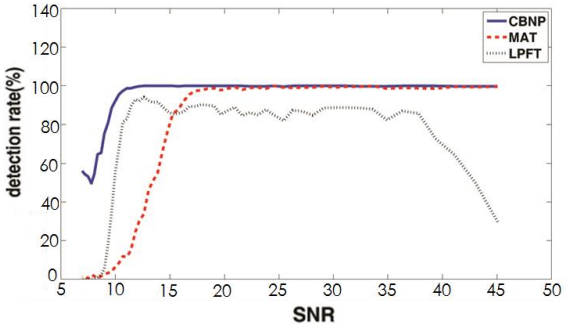


Fig. 18. Deviation rate and SNR

VI. Conclusion and Future Work

In this paper we propose a WSN based vehicle monitoring measurement solution, namely CBNP. Our innovation is that we build vehicle response model for vehicle detection. Instead of traditional energy-threshold method, we perform the vehicle detection task by searching the presence of waveforms similar to the vehicle response model in raw data. To accurately extract vehicle feature, we propose and prove a theorem depicting the relation between feature point and its corresponding point in cross-correlation result. Based on the theorem, we propose a high-quality vehicle feature extraction method. Simulation indicates that our proposed method can outperform the existing methods. In future work, we will focus on inter-node level information fusion, so that the node level information can be used to gain meaningful traffic information.

Appendix A Proof of Theorem

Suppose f is a data sequence which contains a vehicle response c, c is express as $(c_1, c_2, \ldots, c_{N-1}, c_N)$, as we suppose vehicle response to be symmetric, so $c_i = c_{N-i+1}$. To simplifies the case, we suppose the rest part of f is 0. Which means

$$f = (0,0,\ldots,c_1,c_2,\ldots,c_{N-1},c_N,0,0,\ldots),\ c_i = c_{N-i+1} \quad (9)$$

To be general: $f = (a_1, a_2, \ldots)$, correspondingly:

$$a_1 = a_2 = \ldots = a_p = 0 \quad (10)$$

$$a_{p+1} = c_1, a_{p+2} = c_2, \ldots, a_{p+N} = c_N \quad (11)$$

Suppose g is vehicle response model:

$$g = (b_1, b_2, \ldots, b_{M-1}, b_M),\ b_i = b_{M-i+1} \quad (12)$$

Here we suppose both M and N are even. Thus the numerator of cross-correlation at time t is can be calculated as follows.

$$f(t) * g = \sum_{i=1}^{M} (a_{i+t-M} b_{M-i+1}) = \sum_{i=1}^{M/2} (a_{i+t-M} b_{M-i+1} + a_{t-i+1} b_i) \quad (13)$$

According to Cauchy Inequality, we can get the following relationship.

$$a_{i+1-M} b_{m-i+1} + a_{t-i+1} b_i \le \sqrt{(a_{i+t-M}^2 + a_{t-i+1}^2)(b_i^2 + b_{M-i+1}^2)} \quad (14)$$

where "=" happens if and only if $a_{i+t-M} b_{M-i+1} = a_{t-i+1} b_i$.

Thus, we can get the following relationship according to (13) and (14).

$$f(t) * g \le \sum_{i=1}^{M/2} \sqrt{(a_{i+t-M}^2 + a_{t-i+1}^2)(b_i^2 + b_{M-i+1}^2)} \quad (15)$$

where "=" happens if and only if $a_{i+t-M} b_{M-i+1} = a_{t-i+1} b_i$.

Moreover, $\forall i \in [1, M/2], b_i = b_{M-i+1}$, "=" in (15) happens if $\forall i \in [1, M/2], a_{i+t-M} = a_{t-i+1}$.

Furthermore, $\forall i \in [1, N/2]$, when $t = p + \frac{M}{2} + \frac{N}{2}$, we can get the following relationship.

$$c_i = c_{N-i+1}, a_{p+1} = c_1, a_{p+2} = c_2, \ldots, a_{p+N} = c_N \quad (16)$$

$\forall i \in [1, M/2], a_{i+t-M} = a_{t-i+1}$ which means $f(t) * g$ is maximized.

As $t = p + \frac{N}{2}$ the vehicle response peak, so the time delays between cross-correlation peak and vehicle.

References

[1] S. Khan, A. S. K. Pathan, and N. A. Alrajeh, "Wireless sensor networks: Current status and future trends," *Taylor & Francis*, 2013.

[2] R. Zhang, P. Ge, X. Zhou, T. Jiang, and R. Wang, "An method for vehicle-flow detection and tracking in real-time based on Gaussian mixture distribution," *Advances in*

Mechanical Engineering, vol. 13, no. 4, pp. 543-564, Oct. 2013.

[3] R. Wang, L. Zhang, K. Xiao, R. Sun, and L. Cui, "EasiSee: Real-Time vehicle classification and counting via low-cost collaborative sensing," *IEEE Trans. on Intelligent Transportation Systems*, vol. 15, no. 1, pp. 414-424, Feb. 2014.

[4] X. Song, X. Li, W. Tang, *et al*., *A Hybrid Positioning Strategy for Vehicles in a Tunnel Based on RFID and In-Vehicle Sensors*., vol. 14, no. 12, pp. 23095-23118, Dec. 2014.

[5] J. Ding, S. Cheung, C. Tan, and P. Varaiya, "Signal processing of sensor node data for vehicle detection," presented at the IEEE 7th International Intelligent Transportation Systems Conference, Oct. 2004, pp. 70-75.

[6] S. Coleri, S. Cheung, and P. Varaiya, "Sensor networks for monitoring traffic," presented at the Allerton Conference on Communication, Control and Computing, August 2004, pp. 32–40.

[7] D. Li, K. Wong, Y. Hu, and A. Sayeed, "Detection, classification and tracking of targets in distributed sensor networks," *IEEE Signal Processing Magazine*, vol. 2, pp. 17–29, 2002.

[8] L. Gu, D. Jia, P. Vicaire, T. Yan, L. Luo, A. Tirumala, *et al.*, "Lightweight detection and classification for wireless sensor networks in realistic environments," presented at the Proceedings of the 3rd international conference on Embedded Networked Sensor Systems, San Diego, California, Nov. 205-217, 2002.

[9] G. T. Flowers, "Improving reliability of wireless sensor networks for target tracking using wireless acoustic sensors," PhD thesis, Auburn University, 2009.

[10] A. Arora, P. Dutta, S. Bapat, V. Kulathumani, H. Zhang, *et al*., "A line in the sand: A wireless sensor network for target detection, classification and tracking," *Computer Networks*, vol. 46, no. 5, pp. 605–634, July 2004.

[11] X. Wang, M. Zhao, and S. Li, "An improved cross-correlation algorithm based on wavelet transform and energy feature extraction for pipeline leak detection," in *Proc. Better Pipeline Infrastructure for a Better Life*., Wuhan, China, Oct. 2014, pp. 577-591.

[12] M. F. Shinwari, N. Ahmed, H. Humayun, *et al*., "Classification algorithm for feature extraction using linear discriminant analysis and cross-correlation on ECG signals," *International Journal of Advanced Science & Technology*, vol. 48, 2012.

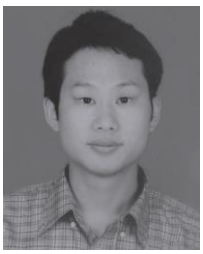
[13] P. Goudard, L. Stehly, F. Brenguier, *et al*., "Cross‐correlation of random fields: Mathematical approach and applications," *Geophysical Prospecting*, vol. 56, no. 3, pp. 375-393, 2008.

[14] S. W. Yelich, *Resistive, Capacitive, Inductive, and Magnetic Sensor Technologies*, CRC Press I Llc, 2014.

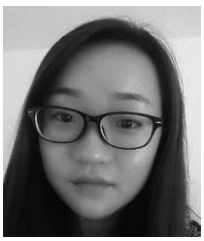
[15] HMC100x. (2015). [Online]. Available: http://www.digikey.com/catalog/en/partgroup/hmc100x-hmc102x-and-hmc1043/12251

Kejiang Xiao received the Ph.D. degree from the Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China, in 2015. He is currently working with the State Grid Information & Communication Company of Hunan Electric Power Company, Changsha, China. His research interests include smart grid, wireless sensor networks, and information fusion and collaboration.

Rui Wang received the Ph.D. degree in pattern recognition and intelligent systems from Northwestern Polytechnical University, Xi'an, China, in 2007. He is currently an Associate Professor with the Department of Computer Science and Technology, School of Computer and Communication Engineering, University of Science and Technology Beijing, Beijing, China. His research interests include wireless sensor networks, intelligent transportation system, self-organization, and information fusion



Lei Zhang received the B.S. degree from East China University of Science and Technology, Shanghai, China, in 2004 and the M.S. degree from the Chinese Academy of Sciences, Beijing, China, in 2011. He is currently working toward the Ph.D. degree at the Department of Computer Science, Georgia State University, Atlanta, GA, USA. His research interests include wireless sensor networks, mobile social networks, and wireless networking.



Lei Xu received the B.S. degree from Henan University of Science and Technology, Henan, China, in 2014. She is currently working toward the M.S. degree from University of Science and Technology Beijing, Beijing, China. Her research interests include wireless sensor networks, and graphical model.