# ROBUST FREEWAY ACCIDENT DETECTION: A TWO–STAGE APPROACH

*Yasitha Warahena Liyanage*[⋆] *Daphney–Stavroula Zois*[⋆] *Charalampos Chelmis*[†]

[⋆]Electrical and Computer Engineering Department
[†]Computer Science Department
University at Albany, SUNY, Albany, NY, USA
Emails: {yliyanage, dzois, cchelmis}@albany.edu

**ABSTRACT**

In this paper, the problem of detecting freeway accidents in real–time based on speed readings from spatially distributed road sensors of variable accuracy is addressed. To ensure robust decision–making, a novel two–stage approach is proposed. Specifically, in the first stage, each sensor generates decisions using a Bayesian quickest change detection framework. In the second stage, individual sensor decisions are aggregated via an optimal stopping approach that optimizes the tradeoff between the costs of aggregation and misclassification. Evaluation of the proposed two–stage approach on a real–world traffic dataset collected from the I405 freeway that passes through the Los Angeles County demonstrates improvements up to 65.2% and 87.2% in average detection delay and probability of false alarm, respectively, as compared to the state–of–the–art.

***Index Terms*—** Intelligent transportation, collision detection, optimal stopping theory, Bayesian quickest change detection, sequential detection

## 1. INTRODUCTION

Traffic accidents in existing transportation networks are characterized by a high fatality rate [1], with both health and cost impacts [2]. Since it may be impossible to prevent traffic accidents from happening altogether, early freeway accident detection has always been desirable. With the emergence of novel sensor network technologies that provide real–time high–fidelity spatiotemporal speed readings about current traffic conditions, real–time accident detection may at last be feasible [3]. However, existing methods either make assumptions about the transportation infrastructure that are currently unsupported, do not explicitly model detection delay and/or false alarm rate, or require knowledge of the relationship between traffic variables and accidents [4–9]. More importantly, the high false alarm rates of the state–of–the–art [8,10], requires robust accident detection schemes to guarantee minimum detection delay without sacrificing detection accuracy.

In our prior work [11,12], we proposed a Bayesian quickest change detection framework to optimally detect accident time in near–real time based on speed sensor readings. Without a proper framework to aggregate individual sensor decisions, however, we relied on naive schemes that (i) assume individual sensor decisions are error–free, (ii) make use of all sensor decisions indiscriminately, and (iii) do not account for the misclassification cost of the final decision. Advanced methods, such as ensemble learning [13, 14], impose restrictions on the base classifiers to operate on the same data stream, and put emphasis on detection accuracy alone, unlike our problem setting, where the time–to–detection is equally important. At the same time, distributed detection [15], where decisions are generated from sensors and subsequently processed at a fusion center to reach consensus, has been well studied [15, 16], with optimum fusion rules obtained under both conditional independent [15] as well as statistically correlated observations [17, 18]. Decision involving unknown signal/noise statistics, and possibly sensors with correlated observations has also been investigated [19–21]. General results about distributed detection under the assumption that neighboring sensors can communicate with each other [22, 23] are also available. In contrast to existing ensemble learning and distributed detection methods, our goal is to derive a framework to maximize freeway accident detection accuracy while minimizing the time–to–detection by optimally integrating sensors' local decisions in an online fashion and at the same time accounting for their individual level of accuracy.

In this work, we address the limitations of state–of–the–art by proposing a two–stage approach that aggregates individual sensor decisions (generated through our Bayesian quickest change detection framework [11, 12]) online based on optimal stopping theory principles. Specifically, at each time step, the proposed framework inspects the accuracy of each sensor, and the costs of aggregation and misclassification of the final decision, to decide whether to make a decision or continue aggregating individual sensor decisions. On real–world speed data, our approach improves average detection delay and probability of false alarm by up to 65.2% and 87.2%, respectively as compared to the state–of–the–art on (i) inferring accident start time [8], and (ii) change point detection in time–series [24], and (iii) our own prior work [11, 12].

## 2. BACKGROUND

We consider a freeway consisting of a set $\mathcal{L}$ of spatially distributed sensors. Each sensor $s_i \in \mathcal{L}$ measures the average speed of passing vehicles to generate a sequence $\{Y_k^i\}$ of speed readings over discrete time indices $k \in \{1, 2, \dots\}$ from which a sequence of accident sensitive features (ASFs) $\{Z_k^i\}$ are derived. At some random time $\nu$, an accident occurs, which changes the distribution of $\{Z_k^i\}$.

In our prior work [11, 12], we have devised a Bayesian quickest change detection framework based on the objectives of average detection delay $d_a(\tau) \triangleq \mathbb{E}\{(\tau - \nu)^+\}$, where $x^+ = \max(0, x)$, and probability of false alarm $P_{\text{FA}}(\tau) \triangleq P(\tau < \nu)$, i.e., probability of false detection of an accident, where $\tau$ denotes the time at which we declare that an accident has happened (referred to as *stopping time* in decision theory [25]), and $\nu$ is the actual accident time. The goal is to select a stopping time $\tau$ on the ASFs sequence $\{Z_k^i\}$, at which to declare an accident, such that $d_a(\tau)$ is minimized subject to a constraint $\gamma \in (0, 1)$ on $P_{\text{FA}}(\tau)$, i.e., $\min_\tau \quad d_a(\tau) \quad \text{s.t.} \quad P_{\text{FA}}(\tau) \leq \gamma$. The solution to this optimization problem is the optimal strategy $\tau_{\text{optimal}} = \inf\left\{k \geq 0 \mid \pi_k^i \geq (1-\gamma)\right\}$, where $\pi_k^i = P(\nu \leqslant k | Z_1^i, \dots, Z_k^i)$ represents the *a posteriori probability* at time $k$ of sensor $s_i$. For implementation convenience, we consider the log–likelihood ratio $g_k^i = \log\left(\frac{\pi_k^i}{1-\pi_k^i}\right)$, a monotone non–decreasing function of $\pi_k^i$, which leads to:

$$\tau_{\text{optimal}} = \inf\left\{k \geq 0 \mid g_k^i \geq \delta^*\right\},\ \delta^* = \log\left(\frac{1-\gamma}{\gamma}\right). \quad (1)$$

In [11, 12], we have shown that the pre– and post–accident distribution of the ASFs sequence is Gaussian, and accident time $\nu$ has a zero modified geometric distribution with parameters $\rho$ and $\pi$, resulting in the following recursive formula:

$$g_k^i = \log(\rho + e^{g_{k-1}^i}) - \log(1-\rho) + \log\left(\frac{\sigma_0}{\sigma_1^2}\right) + \frac{(Z_k^i - \mu_0)^2}{2\sigma_0^2} - \frac{(Z_k^i - \mu_1)^2}{2\sigma_1^2},\ g_0^i = \log\left(\frac{\pi}{1-\pi}\right), \quad (2)$$

where $Z_1^i, Z_2^i, ..., Z_{\nu-1}^i$ are distributed as $\mathcal{N}(\mu_0, \sigma_0^2)$, while $Z_\nu^i, Z_{\nu+1}^i, ...$ are distributed as $\mathcal{N}(\mu_1, \sigma_1^2)$.

## 3. PROBLEM DESCRIPTION

Our goal is to optimally aggregate individual sensor decisions obtained via the log–likelihood ratio test in Eq. (1), to improve the robustness of the accident detection process. Since only a small number of downstream sensors (i.e., direction opposite to the direction of a moving vehicle) within $[0.1, 2]$ miles from an accident location can provide useful information about the accident [11, 12], we restrict the problem of aggregation to a set $\mathcal{L}_N \subset \mathcal{L}$ of $N$ downstream sensors within $d$ miles from the location of each sensor $s_i \in \mathcal{L}$.

We begin by discretizing the decision of each sensor $s_i \in \mathcal{L}_N$ at time $k$ by defining an indicator variable $l_k^i$ as:

$$l_k^i = \begin{cases} 1, & \text{if } g_k^i \geq \delta^* \\ 0, & \text{otherwise} \end{cases}. \quad (3)$$

We model the state $V_k$ of freeway in the vicinity of sensors in $\mathcal{L}_N$ at time $k$ as $V_k = 0$ (*no accident*), or $V_k = 1$ (*accident*), with *a priori probability* $P(V_k = 1) = p, \forall k \in \{1, 2, \dots\}$, where $p$ denotes the probability of an accident. We also consider the accuracy, $h_k^i$, of each $s_i$ at $k$. Intuitively,

$$P(l_k^i = u | V_k = u) = h_k^i,\ \ \forall k \in \{1, 2, \dots\}, \quad (4)$$

where $u \in \{0, 1\}$. Note that $P(l_k^i = 1 | V_k = 1) + P(l_k^i = 0 | V_k = 1) = 1$ and $P(l_k^i = 1 | V_k = 0) + P(l_k^i = 0 | V_k = 0) = 1$. Further, we define misclassification cost $M_{uw} \geqslant 0, u, w \in \{0, 1\}$, as the cost of detecting state $V_k = w$ when the true state is $u$, whereas we use variables $c_i, \forall s_i \in \mathcal{L}_N$, to denote the total cost of decision $l_k^i$, i.e., cost of computing ASF $Z_k^i$, and cost of performing the log–likelihood ratio test in Eq. (1). Note that even if only a small subset of sensors in $\mathcal{L}_N$ pass the log–likelihood ratio test at any time $k$, they will all continue generating $l_k^i$ values (i.e., consecutive ones can be generated) at each subsequent time step until a final decision is made.

In order to detect $V_k, \forall k$, we propose to sequentially examine individual sensor decisions in a decision center with which sensors are connected [26]. At each step, the decision center has to decide between continuing and stopping the aggregation process based on the accumulated information thus far and the cost of evaluating the remaining sensors' decisions. To this end, we introduce a pair of random variables $(R, D_R)$, where $R$ denotes the sensor at which the framework stops at, with the event $\{R = s_i\}$ depending only on the individual decisions of the sensor set $\{s_1, s_2, \dots, s_i\}$, that is the information accumulated up till sensor $s_i$. $D_R$ denotes the decision choice to select among the set $\{0, 1\}$, where the decision $\{D_R = 1\}$ represents declaring an accident based on the information accumulated up until sensor $R$.

Since we have already incorporated the average detection delay and false alarm rate in our initial formulation (see Section 2), what remains is to improve the robustness of the decision–making process by selecting $R$ and $D_R$ at each time $k$ by optimizing the following cost function:

$$J(R, D_R) = \mathbb{E}\left\{\sum_{i=1}^{R} c_i + \sum_{w=0}^{1}\sum_{u=0}^{1} M_{uw} P(D_R = w, V_k = u)\right\}. \quad (5)$$

The first term represents the cost of considering individual sensor decisions, while the second term penalizes the misclassification cost. Thus, our optimization problem reduces to finding a pair of random variables $(R, D_R)$ such that:

$$\underset{R, D_R}{\text{minimize}}\ J(R, D_R). \quad (6)$$

## 4. ROBUST FREEWAY ACCIDENT DETECTION

In order to solve the optimization problem defined in Eq. (6), we consider the *a posteriori probability* $\lambda_i \triangleq P(V_k = 1|l_k^1, \ldots, l_k^i)$, a sufficient statistic of the accumulated information that can be computed recursively as in Lemma 1.

**Lemma 1.** *When sensor $s_i$ generates decision $l_k^i$, the a posteriori probability $\lambda_i$ can be computed as follows:*

$$\lambda_i = \frac{P(l_k^i|V_k = 1)\lambda_{i-1}}{P(l_k^i|V_k = 1)\lambda_{i-1} + P(l_k^i|V_k = 0)(1 - \lambda_{i-1})}, \quad (7)$$

**Lemma 2.** *From the definition of $\lambda_i$ and the fact that $x_R = \sum_{i=1}^{N} x_i \mathbb{1}_{\{R=i\}}$ for any sequence of random variables $\{x_i\}$, where $\mathbb{1}_A$ is the indicator function for event $A$ (i.e., $\mathbb{1}_A = 1$ when $A$ occurs, and $\mathbb{1}_A = 0$ otherwise), the probability $P(D_R = w, V_k = 1)$ can we written as follows:*

$$P(D_R = w, V_k = 1) = \mathbb{E}\left\{\lambda_R \mathbb{1}_{\{D_R=w\}}\right\}. \quad (8)$$

The law of total probability enables us to write:

$$P(D_R = w) = \sum_{u=0}^{1} P(D_R = w, V_k = u). \quad (9)$$

Using Lemma 2, Eq. (9) and the facts that $P(D_R = w, V_k = 0) = \mathbb{E}\left\{(1 - \lambda_R)\mathbb{1}_{\{D_R=w\}}\right\}$, and $P(D_R = w) = \mathbb{E}\left\{\mathbb{1}_{\{D_R=w\}}\right\}$ enables us to rewrite the average cost in Eq. (5) as follows:

$$J(R, D_R) = \mathbb{E}\Bigg\{\sum_{i=1}^{R} c_i + \sum_{w=0}^{1}\Big(M_{0w}(1 - \lambda_R) + M_{1w}\lambda_R\Big) \times \mathbb{1}_{\{D_R=w\}}\Bigg\}. \quad (10)$$

To solve the optimization problem in Eq. (6), we must first obtain the optimal decision $D_R$ for any given stopping time $R$. To this end, we first find a lower bound (independent of $D_R$) for the second term of the cost function in Eq. (10), which is the part of the equation that depends on $D_R$.

**Theorem 3.** *For any decision rule $D_R$ given stopping time $R$, $\sum_{w=0}^{1}(M_{0w}(1 - \lambda_R) + M_{1w}\lambda_R)\mathbb{1}_{\{D_R=w\}} \geqslant g(\lambda_R)$, where $g(\lambda_R) \triangleq \min_{w\in[0,1]}[M_{0w}(1 - \lambda_R) + M_{1w}\lambda_R]$. The optimal rule is defined as follows:*

$$D_R^{optimal} = \arg\min_{w\in[0,1]}[M_{0w}(1 - \lambda_R) + M_{1w}\lambda_R]. \quad (11)$$

From Theorem 3, we conclude that:

$$J(R, D_R) \geqslant J(R, D_R^{optimal}), \text{ where } J(R, D_R^{optimal}) = \min_{D_R} J(R, D_R). \quad (12)$$

Thus, we can reduce the cost function in Eq. (10) to one which depends only on the stopping time $R$ as follows:

$$\widetilde{J}(R) = \mathbb{E}\left\{\sum_{i=1}^{R} c_i + g(\lambda_R)\right\}. \quad (13)$$

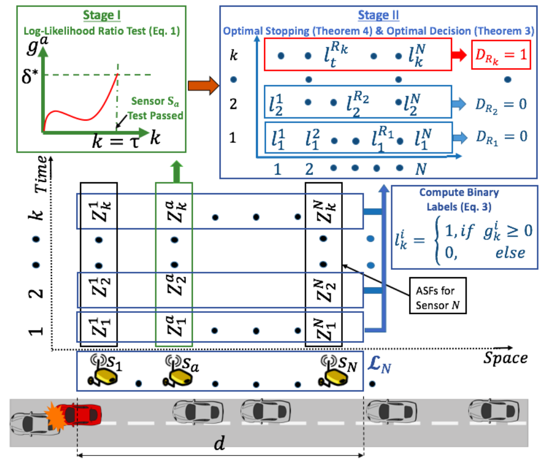


**Fig. 1**. Graphical representation of the proposed approach.

To optimize the cost function in Eq. (13) with respect to $R$, we need to solve the following optimization problem:

$$\min_{R\geqslant 0} \widetilde{J}(R) = \min_{R\geqslant 0} \mathbb{E}\left\{\sum_{i=1}^{R} c_i + g(\lambda_R)\right\}, \quad (14)$$

which constitutes a classical problem in optimal stopping theory for Markov processes [25]. Since $R \in \{0, 1, \ldots, N\}$, the optimum strategy will consist of a maximum of $N + 1$ stages and can be obtained via *dynamic programming* [27].

**Theorem 4.** *For $i = N - 1, \ldots, 0$, the function $\bar{J}_i(\lambda_i)$ is related to $\bar{J}_{i+1}(\lambda_{i+1})$ through the equation:*

$$\bar{J}_i(\lambda_i) = \min\Bigg[g(\lambda_i), c_{i+1} + \sum_{l_k^{i+1}\in[0,1]} A_{i+1}(l_k^{i+1}) \times \bar{J}_{i+1}\left(\frac{P(l_k^{i+1}|V_k = 1)\lambda_i}{A_{i+1}(l_k^{i+1})}\right)\Bigg], \quad (15)$$

*where $A_{i+1}(l_k^{i+1}) = P(l_k^{i+1}|V_k = 1)\lambda_i + P(l_k^{i+1}|V_k = 0)(1 - \lambda_i)$ and $\bar{J}_N(\pi_N) = g(\lambda_N)$.*

The optimal stopping strategy derived from Eq. (15) has a very intuitive structure: it stops at the stage $i$ where the cost of stopping (the first expression) is no greater than the expected cost of continuing given all information accumulated by stage $i$ (the second expression in the minimization).

Fig. 1 shows a graphical representation of the proposed two–stage approach. For a set $\mathcal{L}_N$ of $N$ sensors located within $d$ miles, at time $k = 1$ of stage I, a sequence of ASFs $\{Z_1^i\}, s_i \in \mathcal{L}_N$, are extracted from speed readings $\{Y_1^i\}, s_i \in \mathcal{L}_N$, and $\{g_1^i\}, s_i \in \mathcal{L}_N$, are computed using Eq. (2). Based on the outcome of the individual log–likelihood ratio tests, a set $\{l_1^i\}, s_i \in \mathcal{L}_N$, of decisions are generated using Eq. (3). Next, at time $k = 1$ of stage II, decisions $\{l_1^i\}, s_i \in \mathcal{L}_N$, are evaluated sequentially using Eq. (15) and once the decision to stop aggregating individual sensor decisions at sensor $s_i = R_1$ is reached, the final decision $D_{R_1}$ is generated using Eq. (11).

The proposed two–stage approach operates in an continuous loop, repeating the steps described above $\forall k$ until an accident is detected (i.e., $D_{R_k} = 1$), at which point an alert is raised, all variables are reset to their initial values, and the process of evaluating ASFs resumes. From an implementation point of view, the accuracy $h_i^k$ for each sensor $s_i$, is updated daily [26], whereas efficient computation of a $(N+1) \times \alpha$ matrix via Theorem 4 for different $\alpha$ values of $\lambda_i \in [0, 1]$, where each row corresponds to $N+1$ values $\bar{J}_i(\lambda_i), i = 0, 1, \dots, N$, for set $\mathcal{L}_N$, is achieved by quantizing the interval $[0, 1]$. Since this computation requires only *a priori* information, it can be conducted offline, once daily. Hence, the complexity of calculating $\bar{J}_i(\lambda_i)$ is independent of time.

## 5. EXPERIMENTS

We evaluate our proposed two–stage approach on a real–world dataset comprising $822,049$ speed readings and $1,158$ accident reports collected during the month of October 2013 for a $50$ mile segment of the I405 freeway in the Los Angeles County. Speed readings (measured in mph every $5$ minutes from $6$am to $9$pm everyday) are collected from $223$ sensors placed approximately $0.5$ miles apart in both north and south directions, and correspond to the average speed of all lanes at a specific location. Accident reports (including time, geographical location, type of accident, and direction) are provided by the California Highway Patrol, LA Department of Transportation, and Caltrans. We consider a single ASF, the speed ratio, computed as $Z_k^i = (Y_k^i - \overline{Y_k^i})/\overline{Y_k^i}$, where $\overline{Y_k^i}$ denotes the historical average speed at time $k$ measured by sensor $s_i$. We found the pre–accident mean and variance of the Gaussian distribution to be $-4.85 \times 10^{-5}$ and $9.25 \times 10^{-2}$, respectively. The estimated parameter $\rho$ of the zero modified geometric distribution was $9.1 \times 10^{-3}$ and $\pi$ was set to $0.001$.

Fig. 2 shows the average detection delay for varying probability of false alarm achieved by our proposed two–stage approach. IIG with "Nearest Center" grouping heuristic [8] (used to infer accident start time), methods in [24] (used for change point detection in time–series data), and the best performing schemes from our prior work, i.e., ATTAIN with Weighted Distance (WD) [11] and ATTAIN–ML with Sensor Accuracy (SA) [12] are also included for comparison. Intuitively, as the probability of false alarm increases, the average detection delay decreases. Our two–stage approach outperforms all baselines with respect to average detection delay and probability of false alarm. Specifically, it achieves $87.2\%$ improvement in false alarm rate as compared to IIG ($\lambda = 0.2$), which achieves the lowest average delay among all baselines, and $65.2\%$ improvement in average detection delay as compared to EGADS–OM, that achieves the lowest probability of false alarm among all baselines. Compared to our prior naive aggregation schemes, our two–stage approach improves average detection delay by $32.1\%$ and $15.4\%$ respectively, for the same lowest false alarm rate.

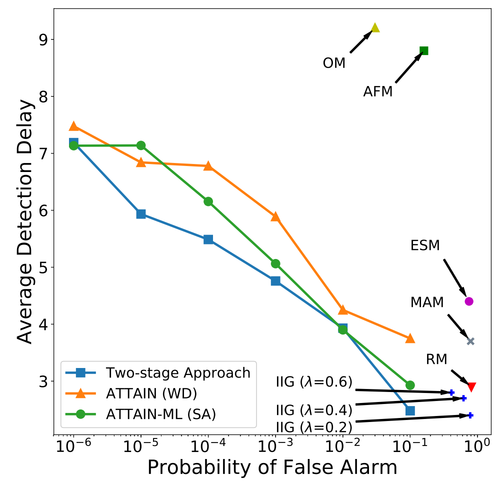


**Fig. 2**. Average detection delay versus false alarm rate. Lines refer to our two–stage approach, ATTAIN (WD) [11] and ATTAIN–ML (SA) [12], crosses to IIG [8] with different thresholds, and rest to models in [24] (i.e., OM: Olympic Model, AFM: Auto Forecast Model, ESM: Exponential Smoothing Model, MAM: Moving Average Model, and RM: Regression Model).

## 6. CONCLUSION

In this work, the problem of robust freeway accident detection was addressed by proposing a two–stage approach that optimally aggregates individual sensor decisions generated by a Bayesian quickest change detection framework. Specifically, an optimization problem was defined in terms of the cost of the aggregation process and the Bayes risk associated with the decision rule, and the optimal solution was derived. The proposed two–stage approach achieves up to $65.2\%$ and $87.2\%$ improvement in the average detection delay and probability of false alarm respectively, compared to the state–of–the–art. Future work will focus on generalizing the proposed framework to jointly detect the accident time and location.

## 7. REFERENCES


[1] National Safety Council, "Nsc motor vehicle fatality estimates," 2017, [Online]. Available: `https://www.nsc.org/Portals/0/Documents/NewsDocuments/2018/December_2017.pdf`.

[2] "WHO Global Status Report On Road Safety 2015," `http://apps.who.int/iris/bitstream/10665/189242/1/9789241565066_eng.pdf?ua=1`.

[3] J. Gubbi, R. Buyya, S. Marusic, and M. Palaniswami, "Internet of Things (IoT): A vision, architectural elements, and future directions," *Future generation computer systems*, vol. 29, no. 7, pp. 1645–1660, 2013.

[4] A. Willsky, E. Chow, S. Gershwin, C. Greene, P. Houpt, and A. Kurkjian, "Dynamic model-based techniques for the detection of incidents on freeways," *IEEE Transactions on Automatic Control*, vol. 25, no. 3, pp. 347–360, 1980.

[5] L. Canzian, U. Demiryurek, and M. Van Der Schaar, "Collision Detection by Networked Sensors," *IEEE Transactions on Signal and Information Processing over Networks*, vol. 2, no. 1, pp. 1–15, March 2016.

[6] S. Yang, K. Kalpakis, and A. Biem, "Detecting road traffic events by coupling multiple timeseries with a nonparametric bayesian method," *IEEE Transactions on Intelligent Transportation Systems*, vol. 15, no. 5, pp. 1936–1946, 2014.

[7] D.-S. Zois, U. Demiryurek, and U. Mitra, "A POMDP approach for active collision detection via networked sensors," in *50th Asilomar Conference on Signals, Systems and Computers*, 2016, pp. 1697–1701.

[8] M. Yue, L. Fan, and C. Shahabi, "Inferring Traffic Incident Start Time with Loop Sensor Data," in *25th ACM International on Conference on Information and Knowledge Management*, 2016, pp. 2481–2484.

[9] K. Zhang and Michael A. P. Taylor, "Effective arterial road incident detection: a Bayesian network based algorithm," *Transportation Research Part C: Emerging Technologies*, vol. 14, no. 6, pp. 403–417, 2006.

[10] M. Ozbayoglu, G. Kucukayan, and E. Dogdu, "A real-time autonomous highway accident detection model based on big data processing and computational intelligence," in *IEEE International Conference on Big Data (Big Data)*, Dec. 2016, pp. 1807–1813.

[11] Y. W. Liyanage, D.-S. Zois, and C. Chelmis, "Optimal Sequential Detection of Freeway Accidents," in *Asilomar Conference on Signals, Systems, and Computers (ACSSC)*, Nov. 2018.

[12] Y. W. Liyanage, D.-S. Zois, and C. Chelmis, "Quickest Freeway Accident Detection Under Unknown Post-Accident Conditions," in *6th IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, Dec. 2018.

[13] T. G. Dietterich, "Ensemble methods in machine learning," in *International workshop on multiple classifier systems*. Springer, 2000, pp. 1–15.

[14] H. Zheng, S. R. Kulkarni, and H. V. Poor, "Attribute-distributed learning: models, limits, and algorithms," *IEEE Transactions on Signal processing*, vol. 59, no. 1, pp. 386–398, 2011.

[15] P. K. Varshney, *Distributed detection and data fusion*, Springer Science & Business Media, 2012.

[16] R. Viswanathan and P. K. Varshney, "Distributed detection with multiple sensors Part I. Fundamentals," *Proceedings of the IEEE*, vol. 85, no. 1, pp. 54–63, 1997.

[17] E. Drakopoulos and C.-C. Lee, "Optimum multisensor fusion of correlated local decisions," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 27, no. 4, pp. 593–606, 1991.

[18] M. Kam, Q. Zhu, and W. S. Gray, "Optimal data fusion of correlated local decisions in multiple sensor detection systems," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 28, no. 3, pp. 916–920, 1992.

[19] P. Willett, P. F. Swaszek, and R. S. Blum, "The good, bad and ugly: distributed detection of a known signal in dependent Gaussian noise," *IEEE Transactions on Signal Processing*, vol. 48, no. 12, pp. 3266–3279, 2000.

[20] B. Chen and P. K. Varshney, "A Bayesian sampling approach to decision fusion using hierarchical models," *IEEE Transactions on signal processing*, vol. 50, no. 8, pp. 1809–1818, 2002.

[21] O. Ozdemir, T. Allen, S. Choi, T. Wimalajeewa, and P. Varshney, "Copula Based Classifier Fusion Under Statistical Dependence," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017.

[22] L. Canzian and M. van der Schaar, "Timely Event Detection by Networked Learners," *IEEE Trans. Signal Processing*, vol. 63, no. 5, pp. 1282–1296, 2015.

[23] N. Katenka, E. Levina, and G. Michailidis, "Local vote decision fusion for target detection in wireless sensor networks," *IEEE Trans. Signal Processing*, vol. 56, no. 1, pp. 329–338, 2008.

[24] N. Laptev, S. Amizadeh, and I. Flint, "Generic and scalable framework for automated time-series anomaly detection," in *21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2015, pp. 1939–1947.

[25] A. N. Shiryaev, *Optimal stopping rules*, vol. 8, Springer Science & Business Media, 2007.

[26] California Department of Transportation, "Performance Measurement System (PeMS)," [Online]. Available: `http://pems.dot.ca.gov/`.

[27] D. P. Bertsekas, *Dynamic Programming and Optimal Control*, vol. 1, Athena Scientific, 2005.