Emergency Incident Detection from Crowdsourced Waze Data using Bayesian Information Fusion

Yasas Senarath
George Mason University
Fairfax, VA, USA
ywijesu@gmu.edu

Saideep Nannapaneni
Wichita State University
Wichita, KS, USA
saideep.nannapaneni@wichita.edu

Hemant Purohit
George Mason University
Fairfax, VA, USA
hpurohit@gmu.edu

Abhishek Dubey
Vanderbilt University
Nashville, TN, USA
abhishek.dubey@vanderbilt.edu

Abstract—The number of emergencies have increased over the years with the growth in urbanization. This pattern has overwhelmed the emergency services with limited resources and demands the optimization of response processes. It is partly due to traditional 'reactive' approach of emergency services to collect data about incidents, where a source initiates a call to the emergency number (e.g., 911 in U.S.), delaying and limiting the potentially optimal response. Crowdsourcing platforms such as Waze provides an opportunity to develop a rapid, 'proactive' approach to collect data about incidents through crowd-generated observational reports. However, the reliability of reporting sources and spatio-temporal uncertainty of the reported incidents challenge the design of such a proactive approach. Thus, this paper presents a novel method for emergency incident detection using noisy crowdsourced Waze data. We propose a principled computational framework based on Bayesian theory to model the uncertainty in the *reliability of crowd-generated reports* and *their integration across space and time* to detect incidents. Extensive experiments using data collected from Waze and the official reported incidents in Nashville, Tenessee in the U.S. show our method can outperform strong baselines for both F1-score and AUC. The application of this work provides an extensible framework to incorporate different noisy data sources for proactive incident detection to improve and optimize emergency response operations in our communities.

Keywords—Emergency Informatics, Bayesian Theory, Information Fusion, User-generated Content, Uncertainty

I. Introduction

Emergency response to incidents such as road accidents is one of the most pressing problems faced by communities across the globe [1]. Given the growing urbanization, this problem is further exaggerated and constrains the limited resources of emergency management agencies. For example, the high dependence on emergency communication lines limits and delays the ability to timely collect data about incidents, especially during disasters. At such times, the damage to centralized emergency operations centers and communication lines often make the collection of information hard and therefore crowdsourced reports, even though sporadic and possibly false, become an important tool (e.g., Hurricane Harvey in 2017 [1]).

There are a number of commercial crowdsourcing platforms. User-generated content on these platforms often provide near real-time observations of the surroundings of the
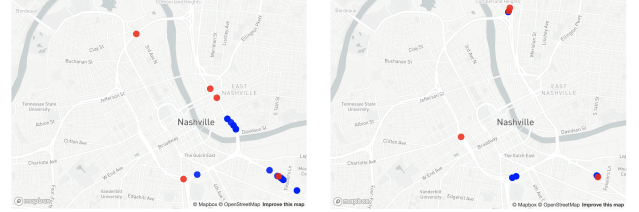

Fig. 1. Figure shows the varied distribution of crowdsourced Waze reports (*blue*) about incidents (*red*) in a region at two different times, illustrating the information fusion challenge for crowdsourced data across space and time.

source user, which can be an invaluable data source in the times of an emergency [2]. However, the inaccuracy of users (location and time of reporting) while creating observational reports about incidents creates challenges, mainly *unreliability of reporting sources* and the *uncertainty of the reported incident information* [3].

If we consider the crowd-sourced report to be a sensor reading and the actual incident requiring emergency response to be a failure, the problem of inferring an incident from a series of sensor readings (crowd-sourced reports) is similar to the classical fault diagnostics problem in cyber-physical systems [4], where the goal is to characterize the system performance using heterogeneous sensor data, and identify any faults. In this paper, we consider the transportation network as the system of interest, and Waze crowdsourcing reports from people (human sensors) as analogous to data from mechanical sensors. Some of the differences between traditional diagnostics problem and the goal of incident detection analysis include: a.) Waze reports from human sensors can be unreliable when compared to mechanical sensors, b.) The data from human sensors can be highly uncertain (e.g. location information), and c.) There is a huge variability across the several Waze reports pertaining to the same incident due to the variability in the human sensor reporting (*c.f.* Figure 1).

Paper Contributions: We present a novel Bayesian-theoretic method for incident detection that systematically models the uncertainty in the spatio-temporal information fusion of noisy crowdsourced reports while considering the ground-truth as conventional reports from the official local government agencies. In this paper, we analyzed Waze Reports from Nashville, Tennessee, USA and considered the incident

information from the Enhanced Tennessee Roadway Information Management System (*E-TRIMS*) as the ground truth.

The proposed approach is novel as it uses a principled Bayesian-theoretic approach for aggregating uncertain crowdsourced data, and for estimating probabilities of occurrence of incidents. The Bayesian approach enables sequential updating of incident probabilities as new Waze reports arrive in time. Also, the quantitative evaluation of incident probabilities provide a measure of confidence of our prediction.

Paper Organization: Section II provides a review of related work. Section III provides a review of information avaialable from incident reports from Waze and E-TRIMS platforms. Section IV details the problem statement of information fusion of Waze data and describes the proposed Bayesian information fusion approach. Section V presents the experimental evaluation study that was carried out to demonstrate the proposed methodology. Section VI presents the results of the experimental case study and a comparison of the proposed method with the existing approaches, and Section VII presents the concluding remarks.

II. BACKGROUND AND RELATED WORK

Transportation systems are the critical infrastructure of the modern cities. Therefore, significant research has been conducted to perform analytics on the transportation systems and traffic conditions using various sources of available data, including crowdsourced data. Hence, we provide a brief literature overview on the use of crowdsourced data for analyzing transportation systems. Silva et al [5] studied the breadth and spatial coverage capabilities of the Waze report information to understand the traffic conditions, and also their limitations. They concluded that the frequency of the Waze alerts is consistent with the users' routines and traffic patterns. Sanchez et al. [6] investigated the impacts of dynamic traffic lighting by analyzing the traffic jam and traffic incident reports from Waze. The authors concluded that dynamic traffic lighting can reduce traffic congestion and improve traffic speed but has no impact on the occurrence of traffic accidents. Eriksson [7] and Santos et al. [8] integrated accident data from the official accident database and Waze to create a comprehensive accident database after removing the data that were common in both the datasets through spatio-temporal proximity analysis.

Lenkei [3] performed a comparative study of the accidents reported through Waze and the official traffic database in Sweden (Trafikverket), and concluded that there was a 43% overlap in the accident data between Waze and Trafikverket. The author provided two reasons for the increased number of Waze alerts. First, some of the Waze alerts can either be self-resolving, false alarms, or of low impact. And, second, the Waze data stream can include multiple alerts referring to the same accident. Lenkei also observed that 27.5% of the incidents in Trafikverket were detected earlier by Waze.

These studies motivate our work to explore the predictability of unconventional, *noisy* Waze reports to detect emergency incidents, and especially, the need for a generic, principled approach for spatio-temporal information fusion of the *redundant* and *noisy* crowdsourced reports to detect emergency incidents.

Furthermore, Flynn et al. [9] investigated the ability of Waze data to serve as a reliable indicator of police-reportable crashes through machine learning approaches. In particular, the authors considered a Random Forest (RF) model to predict a police-reportable crash (binary outcome) using a set of features based on weather, Waze Reports (median report reliability, number of records, type of Waze event, road classification), time (hour of day and day of week), and urban area classification.

In summary, we note that while there exist prior research studies on crowdsourced data for event detection and transportation incidents, there is a lack of a.) a principled approach for information fusion of noisy data to detect incidents that could be extended in general across crowdsourcing platforms because the existing approaches consider only simple aggregation methods (count/average) of surface level features without considering impacts of uncertainty associated with report integration, and b.) a comprehensive analysis of the effect of spatio-temporal resolution in the fusion process of noisy crowdsourced data.

Next, we present our methodology to remedy these gaps in the prior research. First, we provide a brief overview of the Waze and E-TRIMS datasets (with similar trends, *c.f.* Figure 2) used in this paper to train a model for incident prediction.
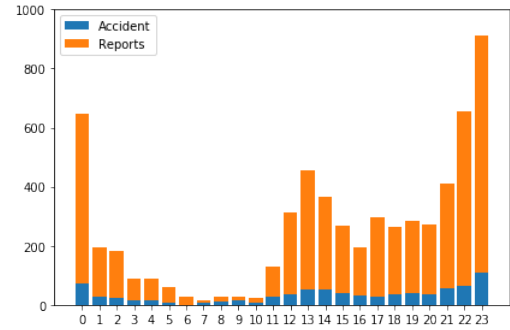

Fig. 2. Similar hourly trend of Waze reporting and accident records in the official E-TRIMS system.

III. DATA PLATFORMS

A. Waze Platform

Waze is a GPS navigation application that enables its users to send reports of different traffic conditions while they are travelling [10]. In the following, we provide the descriptions of most relevant attributes of a user-generated report:

- **type**: Type of traffic alert; this study only uses traffic alerts with type 'ACCIDENT'.
- **confidence**: a score for the report based on other users' reactions to the report, between [0-10].
- **reportRating**: The level of the reporting user credibility, between [1-6].
- **reliability**: Indicates the reliability of the Waze report, where the reliability is based on *confidence* of the report and *reportRating* score of the user; value between [1-10].
- **location**: Geographic coordinates (latitude, longitude) of the report origin.
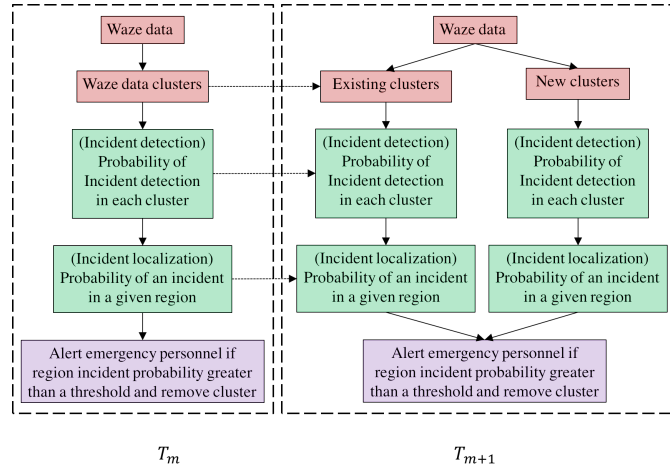- **pubMillis**: Reporting time in milliseconds since epoch.

Fig. 3. Overview of the Bayesian-theoretic approach to fuse information and detect incidents.

B. E-TRIMS Platform

This is a single integrated system in Tennessee state of the U.S. that includes State and local roadways, structures, pavement, traffic, photo logs, and crash data. We use this data source to validate the models for incident detection. In contrast to crowdsourced data from Waze, the E-TRIMS traffic data is more reliable. An incident record in the E-TRIMS constitutes of attributes that are similar to user-generated report on Waze:

- **latitude**: GPS Coordinate Latitude of the E-TRIMS incident record.
- **longitude**: Longitude of the E-TRIMS incident record.
- **timestamp**: Time the accident is recorded in the system in milliseconds since epoch.
- **unit_segment_id**: Id of the road segment that accident has happened.

IV. PROBLEM STATEMENT & PROPOSED METHODOLOGY

A. Definitions and Problem Statement

We summarize the notations used in Table I and define the key concepts as follows:

- **Region** (R_j): We divide the area of interest (W) for incident detection into hexagonal grids called regions, similar to the approach used in [11]. We use H3: Uber's Hexagonal Hierarchical Spatial Index [12]. This enables us to choose a resolution value (Res) and obtain grid indexes corresponding to provided geo-coordinates. Regions are the spatial unit of accident prediction in this study.
- **Incident Time Period** (T'): We assume that an incident in a region R_j will prevail for a T' period. We call this period an incident time period in R_j and assume a single incident occurring in R_j during T'.
- **Time step** (T_m): The smallest time interval at which Waze reports are aggregated ($T_m \in T'$) to detect an incident.

Motivating Example: Let us consider a scenario to understand the problem that is considered in this paper.

Assume that we begin collecting Waze reports at 12:00 am in the intervals of five minutes. Let there be six Waze reports

TABLE I
NOTATION FOR VARIABLES IN METHODOLOGY

Variable	Description
W	Total area of interest (e.g., City of Nashville)
k	Number of non-overlapping regions W is divided into
R_j	j^{th} region
t_s	time interval after which Waze reports are aggregated
T_m	m^{th} time step
$\mathbf{w}_m$	collection of Waze reports available in T_m
h	Number of clusters that the Waze reports are divided into
C_u	u^{th} cluster
$\mathbf{w}_{mu}$	collection of Waze reports available in T_m in C_u
n	Number of Waze reports in $\mathbf{w}_{mu}$
w^i_{mu}	i^{th} Waze report in $\mathbf{w}_{mu}$
l^i_{mu}	latitude value associated with w^i_{mu}
g^i_{mu}	longitude value associated with w^i_{mu}
x^i_{mu}	latitude and longitude information associated with w^i_{mu}, i.e., $x^i_{mu} = [l^i_{mu}, g^i_{mu}]$
R_{ju}	j^{th} region covered by C_u, i.e., the uth cluster
v_u	Number of regions in the area covered by the uth cluster, C_u
r	reliability score associated with a Waze report
p	probability value associated with the reliability score of a Waze report
I_u	binary variable that relates to the occurrence of an incident in the uth cluster. $I_u = 1$ and $I_u = 0$ relates to the occurrence and absence of an incident respectively in the uth cluster
$P(.)$	Probability function
r^i_{mu}	Reliability associated with w^i_{mu}
p^i_{mu}	Probability associated with the reliability of w^i_{mu}
t_r	Reaction time of a reporter between noticing an incident and reporting it on Waze
v_t	Average velocity of a Waze reporter
δ	Distance traveled by a Waze reporter in the duration t_r
$A(w^i_{mu})$	Area around the location of a Waze report, i.e., x^i_{mu} in which an incident may have happened
$\Omega(m,n)$	Function that computes area overlap between any two areas m and n

in the first time step (five minutes) and four Waze reports in the second time step, and similarly, we may have one or more Waze reports in each time step. We would like to use these Waze reports to detect the existence of one or more traffic incidents, and also predict the location of the incident(s).

Given the six Waze reports in the first time step, we will need to infer if these Waze reports are associated with the same traffic incident or multiple incidents. Then, we will need to fuse multiple Waze reports associate with same incident to predict if a traffic incident did occur and also its location.

In the next time step, we have four Waze reports. Now, we will need to decide if these Waze reports are associated with the same incidents identified in the previous time step or if they correspond to new incident(s) not observed in the previous time step. After associating Waze reports to incidents, we will predict the presence of an incident and its location. The subsequent question is to decide when to alert the emergency response authorities about a traffic incident.

Problem Statement: Given a time series of crowdsourced Waze reports, we will need to answer the following questions:

1) Associate available Waze reports to one or more incidents (Sections IV-C and IV-G)
2) Aggregate the Waze reports associated with an incident to estimate the probability of that incident (Section IV-E)
3) Predict the locations of incidents (Section IV-F)
4) Determine when to alert the emergency response authorities about potential incidents (Section IV-H)

B. Overview of the proposed approach

Figure 3 presents an overview of our proposed approach. Since Waze reports are available on continuum in time and space (e.g., city of interest), we implement spatio-temporal discretization and analyze the Waze reports to reduce computational expense and facilitate real-time analysis. After spatio-temporal discretization, we will implement a clustering (grouping) approach on the Waze reports to create Waze data clusters such that all the Waze reports in a cluster are talking about the same incident. The "red" boxes correspond to data gathering and clustering.

We then aggregate the incident information from various Waze reports using a Bayesian information fusion approach to calculate the probability of incident and also calculate the probability of the incident in a given region. The "green" boxes correspond to the incident detection and localization.

If the incident probability in a region was found to be greater than a threshold value, we alert the local emergency response personnel. The threshold value is learned using the Waze reports and incident information available from the local law enforcement agencies. The "purple" box represents decision-making. The threshold value is learned from data by using the E-TRIMS dataset as the ground truth dataset.

If the probability of an incident in any region is less than the threshold, we then consider Waze reports in the next available discrete time step to update the probability of an incident in a region. In this way, we perform sequential updating of incident probabilities. We will also consider a pre-determined number of time steps for updating the incident probabilities. If the incident probability is less than the threshold at the end of all the time steps, we assume an incident did not happen and "delete" the cluster. We explain each of the analysis steps with additional details below.

C. Spatio-Temporal Discretization and Grouping

We perform grouping of several Waze reports into discrete groups where all the reports in a given cluster are assumed to be associated with the same incident. Let T_m represent the mth time step; therefore, $T_m - T_{m-1} = t_s$. Let $\mathbf{w}_m$ represent the set of reports that are available at T_m. Let $C_u, u = 1 \ldots h$ reprsent the h clusters to which $\mathbf{w}_m$ are grouped into. Let $\mathbf{w}_{mu}$ represent the set of n reports in a cluster C_u. We explore two strategies to group Waze reports that identify same incident: segmentation and density-based clustering.

Segmentation Approach: In segmentation method, at the start of incident time period T', we begin to collect Waze reports when the first report appears and stop collecting until T' is over. The process is repeated for each region R_j. We consider all the surrounding regions (R_{ju}) that intersects Waze reports as the region covered by this grouping approach. Advantages of this approach include simplicity and ease of identifying the segments.

Clustering Approach: In addition to following the segmentation process, this approach uses DBSCAN [13] to identify density-based clusters of Waze reports by providing location and time information as features. We explore different values of ϵ parameter in DBSCAN ranging from 0.5 to 0.9, incriminating at 0.1, to find the best set of clusters. For $\epsilon = 0.8$, we obtained the best silhouette score of 0.8156. The resulting clusters are used in our approach to detect incidents. Figure 4 provide the distribution of the cluster time period calculated using Equation 1, where $\mathbf{w}_u$ indicates Waze reports belonging to a cluster C_u. Clusters with zero $TimePeriod$ are excluded in the figures for clarity.

Hereafter the term cluster (C_u) is used to denote the group of Waze reports belonging to the same incident independent of the grouping method employed, unless explicitly mentioned.

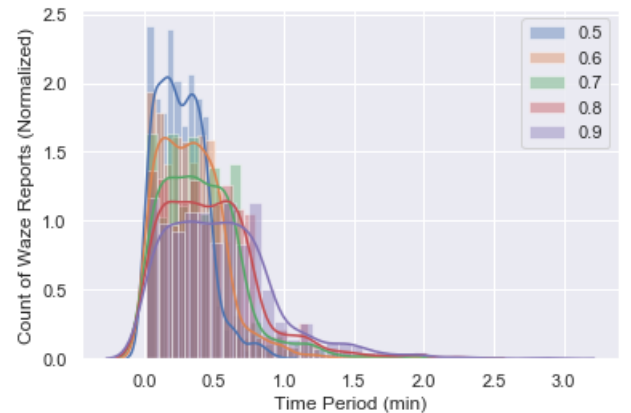$$TimePeriod = max(\mathbf{w}_u.pubMillis) - min(\mathbf{w}_u.pubMillis) \tag{1}$$

Fig. 4. Distribution of time interval (min) of clusters. Colors indicate the value of ϵ.

D. Data Preprocessing

In data preprocessing stage, We calculate prior probability with E-TRIMS data for each hour of the day of each region. Equation 2 illustrates how we calculate the priors for region R_j at hour H of the day. Function $COUNT(region, hour)$ returns the count of E-TRIMS accident records belonging to the provided region and hour of the day from a predefined set of E-TRIMS accident records. $COUNT(*)$ is the number of records in predefined set of E-TRIMS accident records.

$$P(I_u = 1|R_j, H) = \frac{COUNT(R_j, H)}{COUNT(*)} \quad (2)$$

E. Incident Detection

Let W represent the total area of interest for incident detection, and let W be divided into k non-overlapping regions, $R_j, j = 1 \ldots k$. Hence, $R_j \cap R_k = \emptyset, j \neq k$ and $\cup_{j=1}^{k} R_j = W$. These regions can be of different sizes depending on the analysis requirements.

As discussed in Section III-A, each Waze report is associated with location information, i.e., latitude and longitude values. Let w_{mu}^i represent the ith Waze report in the pth cluster at T_m. Let $x_{mu}^i = [l_{mu}^i, g_{mu}^i]$ represent the location of w_{mu}^i.

Let $R_{ju}, j = 1 \ldots v$ represent the v regions that are covered by the Waze reports, and we assume that an incident may have occurred in any of these v regions. Using available Waze reports, we calculate the probability of an incident in each region. Under the assumption that an incident occurred, this probability represents our belief of that incident in each region.

If r is the reliability score ($r \in \mathbb{Z}, 1 \leq r \leq 10$), then the associated probability measure is denoted as p ($0 \leq p \leq 1$). In our current approach, we divide r of each Waze report by ten (max) to obtain the probability p, future work can explore different mechanism to estimate it.

Let $I_u = 1$ and $I_u = 0$ represent the event of an incident happening in any of the v_u regions, and the event of no incident respectively. Let $P(I_u = 1)$ and $P(I_u = 0)$ represent the prior probabilities of an incident and no incident respectively. Let $P(I_u = 1, R_{ju})$ represent the prior probability of an incident in region R_{ju}; therefore, $P(I_u = 1) = \sum_{j=1}^{v_u} P(I_u = 1, R_{ju})$.

Our goal is to obtain the posterior probability of an incident happening in R_{ju} using the available Waze reports $\mathbf{w}_{mu}$; the posterior probability is denoted as $P(I_u = 1, R_{ju}|\mathbf{w}_{mu})$, which can be decomposed into a product of two terms as $P(I_u = 1, R_{ju}|\mathbf{w}_{mu}) = P(I_u = 1|\mathbf{w}_{mu}) \times P(R_{ju}|I_u = 1, \mathbf{w}_{mu})$. The first term, $P(I_u = 1|\mathbf{w}_{mu})$, represents the total probability that an incident happened given a set of Waze reports. Given an incident occurred, the second term, $P(R_{ju}|I_u = 1, \mathbf{w}_{mu})$ computes the probability that it occurred in region R_{ju}.

The probability that an incident occurred or did not occur is calculated using a Naive Bayes approach. For that, we have two classes: (1) Incident occurred ($I_u = 1$), and (2) Incident

did not occur ($I_u = 0$). Using Bayes theorem, the probability of an incident occurrence can be calculated as

$$P(I_u = 1|\mathbf{w}_{mu}) = \frac{P(\mathbf{w}_{mu}|I_u = 1)P(I_u = 1)}{P(\mathbf{w}_{mu})}$$
$$\propto \prod_{i=1}^{n} P(w_{mu}^i|I_u = 1)P(I_u = 1) \quad (3)$$

In Eq. 3, $P(\mathbf{w}_{mu}|I_u = 1)$ represents the likelihood of observing $\mathbf{w}_{mu}$ given that an incident occurred. Assuming all the n Waze reports are independent to each other, the joint likelihood of observing all the n Waze reports can be calculated as the product of the likelihood of the observing individual Waze reports. Similarly, the posterior probability of the absence of an incident can be computed as $P(I_u = 0|\mathbf{w}_{mu}) \propto \prod_{i=1}^{n} P(w_{mu}^i|I_u = 0)P(I_u = 0)$. Therefore, the posterior probability of the occurrence of an incident can be calculated through the normalization, as sum of the two probabilities $P(I_u = 1|\mathbf{w}_{mu})$ and $P(I_u = 0|\mathbf{w}_{mu})$ should add to unity. Therefore,

$$P(I_u = 1|\mathbf{w}_{mu}) = \frac{\prod_{i=1}^{n} P(w_{mu}^i|I_u = 1)P(I_u = 1)}{\sum_{z=0,1} \prod_{i=1}^{n} P(w_{mu}^i|I_u = z)P(I_u = z)} \quad (4)$$

Let p_{mu}^i represents the probability associated with the reliability r_{mu}^i of a Waze report, then the probability that a Waze report corresponds to a true incident is equal to p_{mu}^i (sometimes referred to as true positive probability). Hence, $1 - p_{mu}^i$ can be interpreted as the false positive probability, i.e., probability of a Waze report in the absence of an incident. Eq. 4 can be written as

$$P(I_u = 1|\mathbf{w}_{mu})$$
$$= \frac{\prod_{i=1}^{n} p_{mu}^i P(I_u = 1)}{\prod_{i=1}^{n} p_{mu}^i P(I_u = 1) + \prod_{i=1}^{n} (1 - p_{mu}^i)P(I_u = 0)} \quad (5)$$

F. Incident localization

Typically, after a Wazer identifies an incident, it will take some time for the Wazer to report the incident. First, the user needs to process the incident, decide whether to file a Waze report, and then file a Waze report. The time it takes to perform the above three steps can vary between people; therefore, the location information does not correspond to the true location of the incident, and the incident region is in "close" vicinity to the location in a Waze report. If l_{mu}^i and g_{mu}^i are the latitude and longitude values corresponding to w_{mu}^i, then we consider an area covered by a circle with center (l_{mu}^i, g_{mu}^i) and radius δ. Assume that t_r is the average reaction time of a typical Wazer. If v_t is the velocity of the Wazer, then the distance travelled by a Wazer is $v_t t_r$. We acknowledge that there exists a slight variation of speed limits on various roads/regions, and also variation in the velocities across individual Wazers. In this paper, we do not consider such variations, and assume a constant speed. Therefore, $\delta = v_t t_r$, and is a fixed value.

Let $x_{mu}^i = (l_{mu}^i, g_{mu}^i)$ denote the location of a w_{mu}^i, and $\mathbf{x}_{mu}$ represents the locations of all Waze reports $\mathbf{w}_{mu}$. The

probability of observing $\mathbf{w}_{mu}$ at $\mathbf{x}_{mu}$ given the occurrence of an incident at a region R_{ju} can be calculated as

$$
\begin{aligned}
P(R_{ju}&|I_u = 1, \mathbf{w}_{mu}) \\
&= \frac{P(\mathbf{w}_{mu}|I_u = 1, R_{ju})P(R_{ju}|I_u = 1)}{P(\mathbf{w}_{mu}|I_u = 1)} \\
&\propto P(\mathbf{w}_{mu}|I_u = 1, R_{ju})P(R_{ju}|I_u = 1) \\
&\propto P(\mathbf{w}_{mu}|I_u = 1, R_{ju})P(R_{ju}, I_u = 1)
\end{aligned}
\tag{6}
$$

In Eq. 6, $P(\mathbf{w}_{mu}|I_u = 1, R_{ju})$ is the likelihood of observing $\mathbf{w}_{mu}$ at $\mathbf{x}_{mu}$ given that an incident occurred in R_{ju}. $P(R_{ju}|I_u = 1)$ is the prior conditional probability of an occurrence of an incident in R_{ju} given an incident occurred. As $P(R_{ju}|I_u = 1) = \dfrac{P(R_{ju}, I_u = 1)}{P(I_u = 1)}$, $P(R_{ju}|I_u = 1) \propto P(R_{ju}, I_u = 1)$. $P(R_{ju}|I_u = 1)$ is the prior probability of an incident occurrence in R_{ju} before observing reports $\mathbf{w}_{mu}$.

Assuming independence between the reports, $P(\mathbf{w}_{mu}|I_u = 1, R_{ju})$ is equal to $\prod_{i=1}^{n} P(w_{mu}^i|I_u = 1, R_{ju})$. The probability that a report w_{mu}^i represents an incident in region R_{ju} can be the fraction area overlap between the area covered by w_{mu}^i and R_{ju}. If $A(w_{mu}^i)$ represents the area covered by w_{mu}^i, and $\Omega(m, n)$ the overlap function between any two areas m and n, then the likelihood, $P(w_{mu}^i|I_u = 1, R_{ju})$, is proportional to the area of overlap between the region covered by w_{mu}^i and R_{ju} written as $P(w_{mu}^i|I_u = 1, R_{ju}) \propto \Omega(A(w_{mu}^i), R_{ju})$. Thus, Eq. 6 can be written as

$$
\begin{aligned}
P(R_{ju}&|I_u = 1, \mathbf{w}_{mu}) \\
&\propto P(\mathbf{w}_{mu}|I_u = 1, R_{ju})P(R_{ju}, I_u = 1) \\
&\propto \prod_{i=1}^{n} P(w_{mu}^i|I_u = 1, R_{ju})P(R_{ju}, I_u = 1) \\
&\propto \prod_{i=1}^{n} \Omega(A(w_{mu}^i), R_{ju})P(R_{ju}, I_u = 1)
\end{aligned}
\tag{7}
$$

Given $I_u = 1$, an incident must occur in one of the v regions, $R_{yu}, y = 1 \ldots v$; therefore, $\sum_{y=1}^{v} P(R_{yu}|I_u = 1, \mathbf{w}_{mu}) = 1$. Through normalization, the posterior incident probability given the occurrence of an incident is:

$$
\begin{aligned}
&P(R_{ju}|I_u = 1, \mathbf{w}_{mu}) \\
&= \frac{\prod_{i=1}^{n} \Omega(A(w_{mu}^i), R_{ju})P(R_{ju}, I_u = 1)}{\sum_{y=1}^{v_u} \left(\prod_{i=1}^{n} \Omega(A(w_{mu}^i), R_{yu})P(R_{yu}, I_u = 1) \right)}
\end{aligned}
\tag{8}
$$

The posterior probability of an incident occurring in R_{ju}, i.e., $P(I_u = 1, R_{ju}|\mathbf{w}_{mu})$ can be calculated as a product of $P(I_u = 1|\mathbf{w}_{mu})$ (Eq. 5) and $P(R_{ju}|I_u = 1, \mathbf{w}_{mu})$ (Eq. 8).

G. Waze Report - Incident Association

Section IV-C discussed grouping of Waze reports to segments (clusters) in a time step. As we consider fusing information at discrete time steps, we will have another set of Waze reports in the next step. All the Waze reports in a segment (cluster) is assumed to be associated with an incident.

Let $\mathbf{w}_{m+1}$ represent the set of Waze reports in time step T_{m+1}. Through the Waze Report - Incident Association analysis, we would like to know if a Waze report w_{m+1}^i corresponds to an existing cluster C_u.

In order to realize the above analysis, we first generate a hypothesis that a Waze report belongs to a given cluster, and then test that hypothesis. We use process identified in Section IV-C to associate Waze reports with one of the existing clusters or creating new ones.

After we assign each report to a cluster, we then repeat the analysis in Section IV-E to calculate the updated probabilities of incident detection and localization. It should be noted here the posterior probabilities obtained in the previous time step are used as prior probabilities in the current time step. This use of cluster information and probabilities across two time steps is illustrated using dotted arrows in Figure 3.

H. Incident Classification Model

After calculating the posterior probabilities for a potential incident associated with a region, we use the resulting probabilities as features into a classification model. In this study, we explore different types of binary classification models to identify the optimal value for the probability to create a decision boundary, specifically: Logistic Regression and Random Forest. The classification model identifies the decision boundary to determine whether a given probability value indicates an incident or not. The labels for each prediction time step was obtained by checking the presence of E-TRIMS accident record during the incident time interval associated with the prediction.

We train the classification model using balanced strategy to estimate class weights since the class labels are unbalanced in our datasets. For training classifiers, we use *L-BFGS* algorithm as optimizer for Logistic Regression and the maximum depth of Random Forest Classifier is set to five and three respectively. Other hyperparameters of the models are selected based on a grid search method (details in Section VI.)

V. EXPERIMENTAL EVALUATION

We use five-fold cross validation (CV) setup for evaluation to compare the performance of classification model schemes for incident detection. We used the Waze and E-TRIMS data from October 2019 to December 2019 for validating our schemes. In total there were 33218 Waze reports and 2878 E-TRIMS incident records in the dataset. We analyze the following model schemes using varied sets of features:

- **[M1-M2] Baselines** - Avg. Reliability + Count Features: For each time step t_s, this scheme uses the average score of reliability of Waze reports and the total count of Waze reports in that time step as features.
- **[M3-M4] Proposed** - Plausibility (Clustering-based) Features: This scheme uses our Bayesian-theoretic approach to compute the plausibility score using posterior probabilities of an incident detection in a region. The posterior inference is based on the grouping of Waze reports obtained through clustering approach (*c.f.* Section IV-C). Priors for this scheme is based on E-TRIMS data from Sept., 2019.
- **[M5-M6] Proposed** - Plausibility (Segmentation-based) Features: This scheme is similar to the previous scheme, with one difference of computing the posterior inference for

TABLE II

5-FOLD CV RESULTS FOR THE DIFFERENT INCIDENT DETECTION MODELING SCHEMES. **BOLD FONT** INDICATES THE SCHEMES THAT HAVE MAXIMUM
VALUES FOR F1 AND AUC SCORES.

Model Scheme	Features	Precision	Recall	F1	AUC
[M1, Baseline] Random Forest	Avg. Reliability & Count	31%	78%	44%	64%
[M2, Baseline] Logistic Regression	Avg. Reliability & Count	34%	68%	45%	65%
[M3, Proposed] Random Forest	Plausibility (Clustering-based)	37%	64%	34%	69%
[M4, Proposed] Logistic Regression	Plausibility (Clustering-based)	40%	56%	35%	69%
[M5, Proposed] Random Forest	Plausibility (Segmentation-based)	36%	67%	47%	65%
[M6, Proposed] Logistic Regression	Plausibility (Segmentation-based)	36%	70%	**48%**	66%
[M7, Proposed] Random Forest	All (Clustering-based)	30%	70%	37%	**71%**
[M8, Proposed] Logistic Regression	All (Clustering-based)	30%	66%	36%	**71%**
[M9, Proposed] Random Forest	All (Segmentation-based)	35%	72%	47%	68%
[M10, Proposed] Logistic Regression	All (Segmentation-based)	36%	69%	47%	68%

plausibility scores based on the grouping of Waze reports obtained through segmentation approach (*c.f.* Section IV-C).

- **[M7-M8] All (Clustering-based)** - This scheme uses all of the above features. We used clustering approach for grouping Waze reports for this strategy.
- **[M9-M10] All (Segmentation-based)** - Similar to previous scheme, except we use segmentation approach to group Waze reports.

To evaluate the performance in 5-fold CV, we use the following metrics that are standard for classification models: Precision, Recall, F1 score, and AUC.

VI. RESULT ANALYSIS AND DISCUSSION

This section describes the results for both incident detection models and the impact of spatio-temporal resolution in the modeling. We also present a case study to show the benefits of using crowdsourcing to complement the conventional but slower data collection method to assist emergency response.

After performing grid search on the hyperparameters of the proposed model, we found that hyperparameter values $T' = 25\ mins$, $t_s = 1\ min$, $\delta = 100\ meters$, $Res = 6$ provide the best F1 score. We used those parameters in comparing different model schemes in Section VI-A and case study in Section VI-C. Detailed performance variation of model M6 with different hyperparameter configurations is provided in Section VI-B.

A. Incident Detection

Table II provides the performance of the baselines and proposed methods, where the key observations are the following.

- It is evident that the overall performance of the proposed modeling schemes is better than the baseline approaches in general, based on F1 score and AUC. Moreover, we observe that the results with logistic regression based schemes outperform others.
- The performance of baseline classifier with Avg. Reliability and number of Waze reports as features has better recall (78%) with Random Forest Classifier, however, low precision. Consequently, the F1 score of the baseline classifier is less than the proposed approaches.

- The clustering approach provided better precision at the cost of recall thereby reducing overall F1 score for schemes M3 & M4. It is likely due to dense but accurate clusters that resulted in better precision.

B. Spatio-Temporal Resolution Analysis

Figures 5, 6, and 7 illustrate the variation of different performance scores with respect to four Spatio-Temporal parameters in our model. Figure 6 shows that there is a need to optimize the spatial bound on the unit of analysis for the plausibility score computation for spatial units. Figure 5 reveals that there has been a slight increase in F1 score with increasing incident interval. It is possible that when the incident interval increases there is longer time period to consider it as an accident.
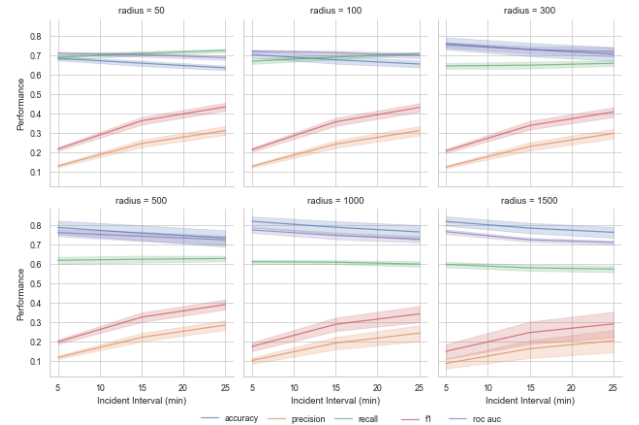

Fig. 5. Variation of performance with incident interval.

C. Case Study

This case study presents the benefits of using our approach for early detection of incidents. To exemplify, consider the incident that happened at $00:05:00$ on $2019-10-01$ (*c.f.* Figure 8). Using our approach, the Waze reports 12 minutes before the E-TRIMS accident record can detect the plausibility of an incident in that region. We collected the time differences between the Waze reports that led to correct accident detection and the corresponding E-TRIMS records. We then averaged
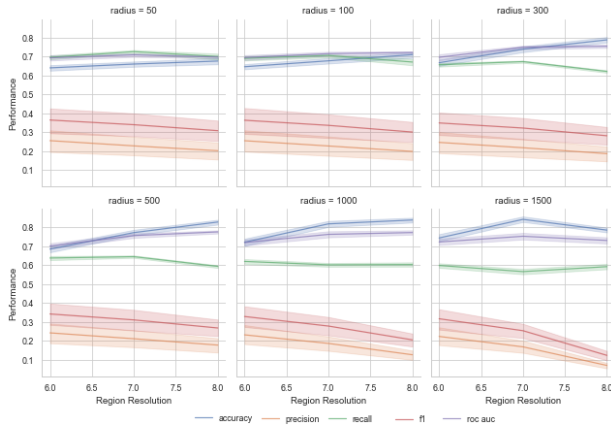
Fig. 6. Variation of performance with aperture size for spatial grids.
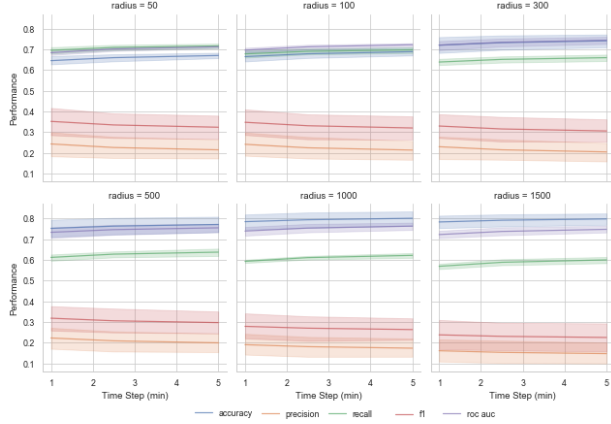

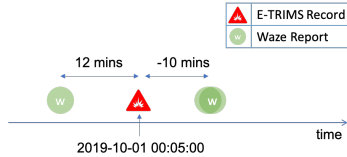Fig. 7. Variation of performance with time step.

Fig. 8. Case study scenario for timing in accident detection in a region.

the detection times to get an average accident prediction time. Any prediction that comes before the official record from E-TRIMS is considered to have a positive prediction time and a prediction that comes afterwards is considered to have a negative prediction time. We found that the average accident prediction time of our best model (Model Scheme - M6) is $5.92\ minutes$. This evidence suggests our approach can help build predictive systems for incident detection using unconventional crowdsourced data, before a conventional system.

VII. Conclusion

This paper presents a principled methodology based on Bayesian theory to address key challenges of modeling the uncertainty and unreliability of crowd-generated reports in their integration across space and time, to detect emergency incidents earlier than the official reporting mechanisms. Our

experiments using Waze data and official reported incidents in Nashville region validate our method with relative gain in F1-score over 5% against baselines. The application of this research can help emergency response operations in systematically incorporating unconventional crowdsourcing data.

Limitations and Future Work. We note the limitation in accurately matching crowdsourced data with the formal incident reports, given the human errors and systematic delays in emergency communication. We can build upon this study to explore automated feature extraction and fusion models to improve the performance, and also, improve the fidelity of the analysis (by moving from regional grids to road segments).

Reproducibility. The code for experiments is available at https://github.com/ysenarath/emergency-incident-detection-web-intelligence-2020.

References

[1] A. Mukhopadhyay, G. Pettet, S. Vazirizade, Y. Vorobeychik, M. Kochenderfer, and A. Dubey, "A review of emergency incident prediction, resource allocation and dispatch models," *arXiv preprint arXiv:2006.04200*, 2020.

[2] M. Amin-Naseri, P. Chakraborty, A. Sharma, S. B. Gilbert, and M. Hong, "Evaluating the reliability, coverage, and added value of crowdsourced traffic incident reports from waze," *Transportation research record*, vol. 2672, no. 43, pp. 34–43, 2018.

[3] Z. Lenkei, "Crowdsourced traffic information in traffic management: Evaluation of traffic information from waze," 2018.

[4] S. Abdelwahed, A. Dubey, G. Karsai, and N. Mahadevan, "Model-based tools and techniques for real-time system and software health management," *Machine Learning and Knowledge Discovery for Engineering Systems Health Management*, vol. 285, 2011.

[5] T. H. Silva, P. O. V. De Melo, A. C. Viana, J. M. Almeida, J. Salles, and A. A. Loureiro, "Traffic condition is more than colored lines on a map: characterization of waze alerts," in *International Conference on Social Informatics*. Springer, 2013, pp. 309–318.

[6] R. Sanchez, D. Martinez, O. A. Mitnik, P. Yanez-Pagans, M. L. Lanzalot, R. Stucchi, and L. Sanguino, "Dynamic traffic lights and urban mobility: An application of waze data to the city of medellın," 2019.

[7] I. Eriksson, "Towards integrating crowdsourced and official traffic data: A study on the integration of data from waze in traffic management in stockholm, sweden," 2019.

[8] S. R. dos Santos, C. A. Davis Jr, and R. Smarzaro, "Analyzing traffic accidents based on the integration of official and crowdsourced data," *Journal of Information and Data Management*, vol. 8, no. 1, pp. 67–67, 2017.

[9] D. F. Flynn, M. M. Gilmore, E. A. Sudderth *et al.*, "Estimating traffic crash counts using crowdsourced data: Pilot analysis of 2017 waze data and police accident reports in maryland," John A. Volpe National Transportation Systems Center (US), Tech. Rep., 2018.

[10] Waze, "Driving Directions, Traffic Reports & Carpool Rideshares by Waze." [Online]. Available: https://www.waze.com/

[11] A. Mukhopadhyay, G. Pettet, C. Samal, A. Dubey, and Y. Vorobeychik, "An online decision-theoretic pipeline for responder dispatch," in *Proceedings of the 10th ACM/IEEE International Conference on Cyber-Physical Systems*, 2019, pp. 185–196.

[12] Uber, "H3: Uber's hexagonal hierarchical spatial index," Aug. 2020, publisher: Uber. [Online]. Available: https://h3geo.org/docs

[13] M. Ester, H.-P. Kriegel, J. Sander, X. Xu *et al.*, "A density-based algorithm for discovering clusters in large spatial databases with noise," in *KDD*, vol. 96, no. 34, 1996, pp. 226–231.