

Detection of Mass Panic using Internet of Things and Machine Learning

Gehan Yahya Alsalat, Mohammad El-Ramly, Aly Aly Fahmy
Faculty of Computer and Information
Computer Science Department, Cairo University
Cairo, Egypt

Karim Said, M.D.
Faculty of Medicine,
Cardiovascular Department
Cairo University
Cairo, Egypt

Abstract—The increase of emergency situations that cause mass panic in mass gatherings, such as terrorist attacks, random shooting, stampede, and fires, sheds light on the fact that advancements in technology should contribute in timely detecting and reporting serious crowd abnormal behaviour. The new paradigm of the ‘Internet of Things’ (IoT) can contribute to that. In this study, a method for real-time detection of abnormal crowd behaviour in mass gatherings is proposed. This system is based on advanced wireless connections, wearable sensors and machine learning technologies. It is a new crowdsourcing approach that considers humans themselves as the surveillance devices that exist everywhere. A sufficient number of the event’s attendees are supposed to wear an electronic wristband which contains a heart rate sensor, motion sensors and an assisted-GPS, and has a wireless connection. It detects the abnormal behaviour by detecting heart rate increase and abnormal motion. Due to the unavailability of public bio-dataset on mass panic, dataset of this study was collected from 89 subjects wearing the above-mentioned wristband and generating 1054 data samples. Two types of data collected were: firstly, the data of normal daily activities and secondly, the data of abnormal activities resembling the behaviour of escape panic. Moreover, another abnormal dataset was synthetically generated to simulate panic with limited motion. In our proposed approach, two-phases of data analysis are done. Phase-I is a deep machine learning model that was used to analyze the sensors’ collected readings of the wristband and detect if the person has indeed panicked in order to send alerting signals. While phase-II data analysis takes place in the monitoring server that receives the alerting signals to conclude if it is a mass panic incident or a false positive case. Our experiments demonstrate that the proposed system can offer a reliable, accurate, and fast solution for panic detection. This experiment uses the Hajj pilgrimage as a case study.

Keywords—Internet of Things; IoT; Mobile Crowd Sensing (MCS); wearables; mass panic; mass gatherings; accelerometer; Optical Heart Rate (HR) sensor; abnormal crowd behaviour; deep learning; Recurrent Neural Network (RNN); Long Short Term Memory (LSTM); Gated Recurrent Unit (GRU); time series

I. INTRODUCTION

The increase of emergency situations cause mass panic, such as terrorist attacks, random shooting, stampede, natural catastrophes, and fires requires fast detection and swift action to save lives. Advances in technology can greatly contribute to the timely detection and reporting of serious crowd abnormal behaviour. Early detection of an incident will give the

emergency authorities valuable time to deal with the situation and prevent it from getting worse by implementing immediate and possibly automated actions.

Mass gathering is an event involving the gathering of a large number of people, at least 1000 but can rise up to millions, at a specific location for a defined period of time and for a specific purpose, it can be either organized or unplanned [1]. Mass gatherings are exposed to unpleasant incidents due to the large number of people and limited space and exit routes. Examples of mass gatherings are major sporting, religious, and cultural events (e.g. Olympic Games, religious pilgrimages, etc.).

A. Mobile Crowd Sensing (MCS)

Mobile Crowd Sensing (MCS) is a new sensing paradigm based on collecting real-time data from two participatory sources: sensing and social media platforms. Therefore, it allows ordinary users to contribute by sharing real-time data with data sensed or collected from their mobile devices/wearables. MCS collects data from users’ devices to analyze them and identify spatiotemporal patterns [2].

The revolution of cost-effective hardware, emergent computing and communication trends such as the Internet of Things (IoT), big data, machine learning, wearable sensing and cloud computing have enabled the existence of mobile crowd sensing applications and made our environment smarter.

B. Internet of Things

Internet of Things (IoT), also known as machine-to-machine (M2M) is a paradigm in which smart sensors and devices collect data and interact with one another without human intervention. Gartner expects that “8.4 billion connected things will be in use worldwide in 2017, up 31 percent from 2016, and will reach 20.4 billion by 2020” [3]. Owing to the rapid growth and advancement in the field of IoT wearable sensors, it became possible to monitor physiological signals continuously, accurately and in a real-time manner [4].

C. Problem Statement and Motivation

Despite the best efforts of authorities to secure mass gatherings, unfortunate incidents still occur in such occasions and cause loss of lives. The causalities of any disaster may not be prevented. However, the early detection in a timely manner

may enable swift actions and reduce the losses. The minutes or even the seconds can contribute to saving people lives.

The main motivation for our work is to overcome the limitations of the existing approaches of detecting mass panic such as video surveillance, audio surveillance, and human surveillance. The details of their shortcomings are explained in Section III. So our proposed system will leverage the popularity of smart wearables and the power of crowd-sourcing, participatory sensing to enable a new wave of effective techniques to detect and report incidents in a timely manner.

D. The Contribution of this Paper

- This paper is – to our best knowledge – the first proposal that employs human physiological bio-data (heart rate and accelerometer) in a mobile crowd sensing (MCS) and IoT application; to contribute to real-time mass panic detection in a mass gatherings.
- We believe that, our dataset collection and data preprocessing is an important contribution that can be a good seed for similar studies. Due to ethical and impractical limitations restricting collecting a real mass panic dataset, it is inevitable to produce artificial datasets reflecting the typical panic behaviour. We plan to disclose our dataset for benchmarking and to make it available for researchers.
- Data analysis phase-I in the wristband: Detecting the panic of an individual was done by analyzing his/her heart rate with respective motion. This problem is a time series problem. We used promising sequential deep machine learning models; Recurrent Neural Networks (RNN), particularly Long Short-Term Memory (LSTM) and Gated Recurrent Units (GRU) as time series classification detection algorithm and we compared their performance with non-sequential models, Vanilla Neural Network (VNN) and Support Vector Machine (SVM).
- Data analysis phase-II: Currently, we propose a simple data analysis module that works on the monitoring server. It can monitor and cluster the incoming signals and differentiates between real mass panic and false positive. This proposed module has data visualization capabilities which can provide event's officials with a real-time view of one or more critical incidents at the same screen. This cannot be achieved using conventional surveillance approaches. Therefore, officials can quickly grasp the threats and take proper fast actions.

E. Paper Scope

It is important to emphasize the following: (1) the scope of this paper is mainly proposing a new detection method based on MCS and IoT, collecting and analyzing experimental data, and running a proof of concept experiment to validate this proposed approach. (2) We did not cover or experiment the details of various possible types of communication networks that may be used to send/receive the sensor data. (3) In this study, we focused on incidents that happen in rapid onset

disasters and occur swiftly in a mass gathering within seconds or minutes and often without any warning. (e.g., terrorist attacks, random shooting, stampede, and fires). (4) Data analysis phase-II prototype is under development and it is planned in future work.

The remainder of this paper is organized as follows. Section II describes the characteristics of mass panic. Section III describes the related work of detecting mass panic. In Section IV, we explain our proposed system architecture and modules. Section V is the case study that includes the experiments and results. Section VI is the discussion section. We conclude the paper in Section VII, and finally, recommend future work.

II. CHARACTERISTICS OF MASS PANIC

Mass Panic is “type of anomaly in a human crowd, which appears when a group of people start to move faster than the usual speed. Such situations can arise due to a fearsome activity near a crowd such as stampede, fire, fight, robbery, riot, etc.” [5]. Here we describe the characteristics of mass panic. Studying such characteristics helps developing detection approaches that exploit technologies which can discover these characteristics. These characteristics can be categorized or explored from two perspectives; first, from physical behaviour of the mass people and second, from physiological factors.

A. The Physical Characteristics of Mass Panic

The physical behaviour characteristics are reflecting the fight and flight response and they are highlighted by Helbing *et al.* and by Bracha [6], [7]. They include (1) People move faster than normal. (2) Individuals start pushing and interactions among people become physical in nature. (3) The physical interactions in the jammed crowd cause dangerous pressures up to 4,450 Nm⁻¹ (refs which can bend steel barriers or push down brick walls). (4) People show a tendency towards mass behaviour, that is, to do what other people do. (5) Many studies [8]-[10] has shown that during disasters, humans evolved to fight, flight, or freeze; the phrase “fight or flight” normally describe the key behaviours that are triggered by fear and occur as a defensive reaction to threat. The temporal relations are summed up by Bracha [7] in terms of “freeze-flight-fight-fright-faint” and he has ordered these sympathetic responses as they occur. A freeze response is believed to be the first response to a threat by some individuals. freeze response was mentioned by Leach [9] as the “stop, look, and listen” to assess a dangerous situation and sometimes it is called “play dead.”. Fight response is when acting impulsively and combat to prevent something undesirable. Flight response is escaping very fast toward a safe place or to run away from serious danger.

B. The Psychological Characteristics of Mass Panic

Here we discuss the characteristic of escape panic from psychological factors. It is mainly about several biological changes in our body that are mediated by the autonomic nervous system. During severe fear, there are significant changes in our physiological measures. For example, the heart rate increases, our breathing becomes rapid, our muscles tense[11]. The sympathetic nervous of the body stimulates the

adrenal glands, which in turn trigger the release of adrenaline and nor-adrenaline hormones. This is what causes the body to respond to stress or fear and as a trigger for the fight-or-flight response which eventually increases the heart rate, breathing rate and movement[12].

III. RELATED WORK

A. General Approaches of Mass Panic Detection

Firstly, we reviewed the possible existing and emergent techniques for mass panic detection in mass gatherings. We classified them into two categories; first, surveillance based techniques; Second, mobile crowd sensing-based techniques.

1) Surveillance based Techniques

Someone or something is monitoring the crowd in the event and reporting any abnormal behaviour to event's authorities. Examples of these approaches are video surveillance, audio surveillance, and human surveillance.

These three approaches have common limitations such as; they are mostly suitable for limited spaces and they cannot cover the whole space for either cost or privacy concerns. Furthermore, each one of these approaches has limitations in the detection method itself. In the video surveillance: (1) If the cameras are smart, they may fail to detect the abnormality in a high dense crowd due to ambiguities and severe occlusions. While if the cameras are traditional; they require human operators to monitor the surveillance cameras continuously over a long period of time, human are subjected to fatigue and loss their attention. (2) The cameras are almost useless during night-time, due to sudden illumination changes. (3) The high computational requirements of the detection algorithms to work in real life [13], [14]. In Audio surveillance, some events may have sounds of very high volume. This makes detection of abnormal sounds difficult [15]. In Human surveillance, the public safety officers are humans that are subject to the same dangers that cause mass panic. They themselves might be victims of the attack or incident that took place or they might panic out of fear for their lives and forget/fail to report the incident.

2) Mobile Crowd Sensing based Techniques

This technique depends on participatory crowd sensing; when normal people who are attending an event contribute to detecting and reporting incidents using their own mobile phones in a timely manner. The Advantages of this approach is that it is cheap and it takes an advantage of attendees' resources (e.g. Mobile phones) and availability. Examples of this approaches are: (1) Contacting emergency authorities by calling 911 [16], [17]. (2) Using social media to report the incident (e.g. posting in Twitter)[18-20]. The limitations of these approaches are: (1) They may not provide real-time detection but near real-time, usually within minutes of the incident, this depends on how fast people will report the incident. (2) Due to the awe and shock of the incident, people might be too busy trying to escape and save their lives. Also, it is possible that they lose their mobile phones during the hide or escape.

It is important to emphasize that, our proposed system is categorized in this approach. The key features of our proposed

system over the previously mentioned approaches are the followings: (1) it provides real-time and fast detection because our technique does not require deliberate action from the user at the time of the incident; the detection and the reporting are done automatically. (2) It covers the whole event premises and it has a good visualization capabilities, where a threat or more than one can be detected and visualized in one screen.

B. Biosensors based Techniques

Secondly, we reviewed the studies that use bio-sensors to detect escape panic. It is important to highlight that – to our best knowledge – there is no real deployment of a large-scale bio-crowd sensing for mass panic detection. But there are a scarce number of proposals or researchers that propose a method of detecting abnormal behaviours of individuals or a small number of people in small buildings.

The U.S. Patent (No. 8,477,035) [21], is a proposed security system triggered by heart rate detection; it is proposed for banks and other enterprises which use security systems in their facilities, to detect robbery or other types of threatening activities. The method includes a wearable heart rate sensor that monitors the heart rate of an individual (security guard or employee); determining if the detected heart rate is abnormal; generating a signal which is wirelessly transmitted to the central monitoring station. The detection is based on threshold-based technique.

The U.S. Patent (No. 9,858,794) [22] is a proposed system for checking the state of workers who are performing tasks in a hazardous environment. The proposed sensing device includes a plurality of biometric sensors to report the corresponding hazardous state of a person. The detection is based on threshold-based technique.

Using accelerometer data, a system [23] with the name Emergency Rescue Support System (ERESS) proposed a disaster detection algorithm. It uses plural sensors such as acceleration, angular velocity, and geomagnetism using data from sensors mounted on ERESS terminals such as smartphones or tablets. ERESS detects the disaster from the behaviour analysis of people by the sensors using machine learning based on group learning using support vector domain description (SVDD). ERESS operates under mobile ad-hoc networks (MANET) between the hand-held terminals.

Using electro-dermal activity and acceleration (skin conductance) the work [24] proposed an algorithm which detects people's extraordinary condition using Skin Potential Level (SPL) sensor (SPN-01) which can measure Skin Potential Level (SPL) with a 9 axis motion sensor to detect an extraordinary condition in case of panic. An electro-dermal activity is a parameter which shows man's mental activity condition and they are electrical signals that appear on an outer skin and sweat glands.

The key differences between our proposed system and the previously mentioned studies [21]-[24] are the followings: (1) They use threshold-based approaches which were designed to detect abnormality of individuals and were not designed for the mass gatherings where knowing the abnormal threshold of a huge number of attendees in advance is challenging. (2) They are designed for small buildings compared to our

proposed system which is intended to cover a very large space using long-range wireless networks. (3) Our proposed system is fusing heart rate sensor with accelerometer sensor. (4) We employed state-of-the-art deep learning models, namely, (Recurrent Neural Network RNNs) and compare them with conventional machine learning approaches.

IV. PROPOSED SYSTEM BASED ON IOT

Our proposed system consists of three modules as follows:

A. Data Acquisition and Data Analysis Phase-I Module

A sufficient number of event attendees are supposed to wear the electronic waterproof wristband. This wristband contains HR and motion sensors, Also it will be embedded by an A-GPS (Assisted-Global Positioning System) for finding the location of a person (by considering the longitude and latitude). The HR and motion sensors are continuously monitoring the heart rate and movement of a person. Once a panic is detected, an alerting signal will be transmitted to the central monitoring server containing the following: User Unique Identifier (UID), location coordinates, heart rate reading, motion data, and timestamp. The monitoring server will distinguish whether the coming signals are a false positive or a real panic incident and then will report the incident severity.

Our assumption is that the wristband will sense the user's heart rate and motion passively, (not connecting to the monitoring server for 24/7) and it will start a connection and send alerting signals once the panic is detected. We have built a machine learning detection model that classifies sensors data as normal or abnormal. Fig. 1 is illustrating the flow chart of the data analysis that occurs in the wristband device. The system modules and the components of the wristband are illustrated in Fig. 2.

B. Communication Module

Every wristband will be embedded with communication modem for wireless connection. The proposed system can use the existing cellular network infrastructure for the wireless communication-between the wristband and the monitoring server- for the following reasons: (1) Service providers have already covered the event area with such networks (GSM/GPRS, 2G, 3.5G or 4G); no need for extra cost of deploying other types of wireless sensors network (WSN). (2) Cellular networks are long-range wirelesses which are able to cover almost all event area.

C. Data Analysis Phase-II and Visualisation Module

In this module, data analysis phase-II is handled by the monitoring server that receives the signals coming from the attendees' wristbands. According to our assumption that was mentioned previously, this monitoring server will receive only the alerting signals. If there is a massive number of alerting signals from the same location, a cautionary warning will be issued and a pop-up colorful pinpoint will appear in the event map at that particular location, with a counter that shows the number of alerting signals per second and this number may increase or decrease depending on the development of the situation.

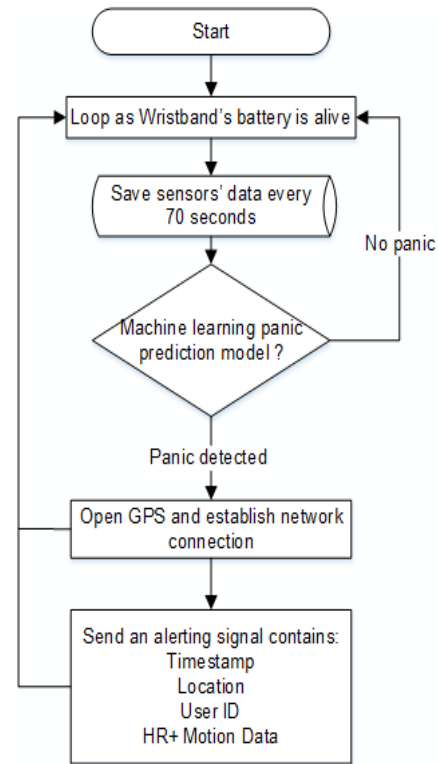


Fig. 1. Flow chart of the data analysis phase-I module.

The color of the pinpoint indicates the severity of the problem; the green color means no threat, it may be a “false positive” alarm, while the yellow color means a caution that should be considered, as there is a probability of having a serious problem. However, when the color goes red, this means a very “critical problem” is going on. Every color represents a range of numbers depending on the number of the event's attendees and the location depth and width. These analytics and data visualization capabilities provide event's officials with a real-time view of critical incidents on one screen and therefore they can take proper action and fast response (see Fig. 3).

V. CASE STUDY: PANIC DETECTION DURING HAJJ PILGRIMAGE

Hajj is the fifth pillar in Islam, It is an annual religious duty which lasts for six days, and it must be performed at least once in a Muslim's lifetime, only for those who are physically and financially capable to afford it [25]. It is one of the most congested mass gatherings in the world. Millions of pilgrims are performing the Hajj rituals in Saudi Arabia and they often crowd into packed spaces at densities of about six or seven persons per square meter [26]. Such crowd densities present a huge challenge for both public safety authorities and any computer-aided system to predict and detect crowd disasters. Therefore we chose Hajj pilgrimage as our case study. In the following, we described: (1) Hardware details. (2) Data collection and preprocessing. (3) Machine learning models. (4) Result and performance metrics.

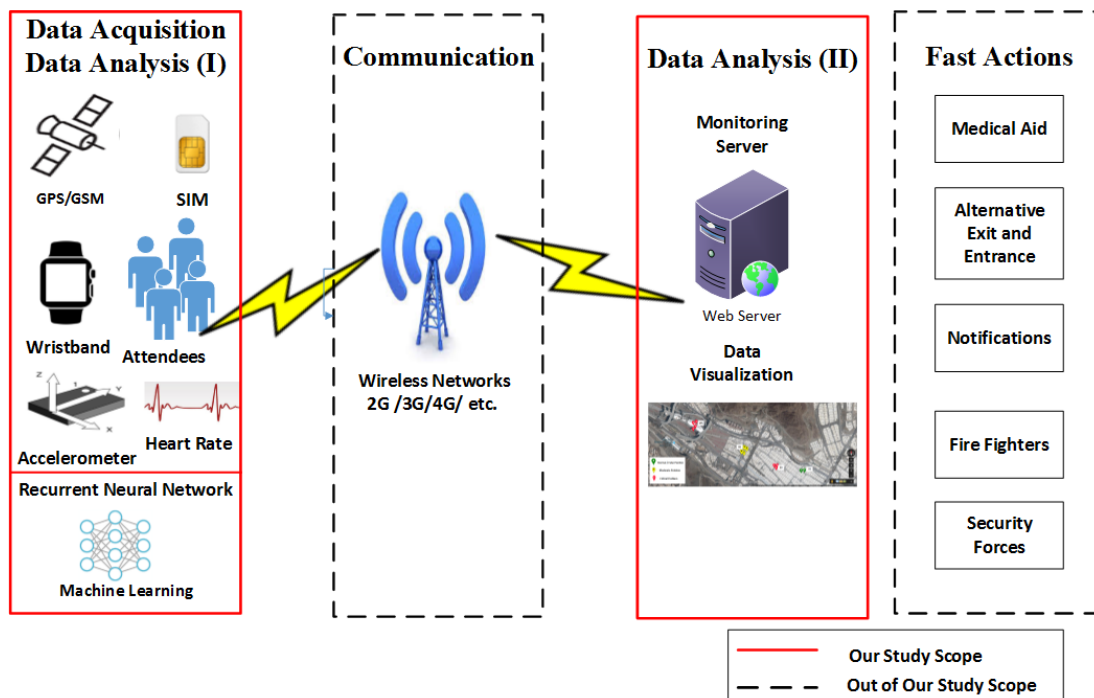


Fig. 2. Proposed system components and modules.

A. Hardware Details

TomTom Spark Cardio wristband [27] was used in this study as our bio-data collection device (see Fig. 4). It is a waterproof wristband that has a built-in optical heart rate sensor that measures the heart rate beats per minute (bpm) using Photoplethysmography (PPG) optical technology. The heartbeat is measured by using a light that measures changes in the blood flow. This is done on the top of the wrist by shining light through the skin. The PPG sensor monitors changes in the light intensity via reflection from or transmission through the tissue and detecting the changing light reflections [28]. Furthermore, it also has an x-axis accelerometer and a Gyro-meter. Moreover, it records the heart rate value with respective acceleratory data in one second interval time. The data is extracted as a csv file.

B. Data Collection

Unfortunately, there exists no publicly available dataset on mass panic as a bio-data. Producing such real dataset is not possible in the real world since it often requires exposing real people to the actual, possibly dangerous environment. These pose ethical and safety concerns. Thus we proposed in this paper, a way to produce an artificial/pretended data on

abnormal activities reflecting on typical behaviour of escape panic. Two types of data were collected, firstly the data of normal daily activities, and secondly the data of abnormal activities (subjects pretended a typical behaviour of escape panic) and another abnormal data was synthetically generated.

C. Subjects and Activities Details

The data collection and subjects' selection were supervised by the fourth author. We consulted him on the physiological factors that cause heart rate fluctuation and also about the data collection activity types, to ensure that we cover all possible normal and abnormal activities that reflect the typical normal and abnormal activities during Hajj rituals. This ensures that the machine learning will learn all possible scenarios to reduce false negative and false positive rates. We conducted our experiments on 89 individuals that are fit, healthy and medication free, aged between 20 and 48 years. This age group represents a good portion of human-beings. Therefore, in our proposed system we targeted this age group (fit and medication free) to be as human sensors for the following reasons: (1) People in this age group are proactive and have the ability to respond to danger and move faster. (2) They are medication free, so there are no other factors that increase HR rather than the real panic situation.



Fig. 3. Data visualization, data analysis phase-II module.



Fig. 4. Tom tom spark cardio wristband.

All participants were informed about the experimental setup and they agreed to contribute to this study. Subjects were asked to perform the following activities: (1) To pretend the escape panic by doing abnormal actions (e.g. running suddenly and rapidly); their heart rate and respective movement data are recorded in one second interval time. It is important to emphasize that, from a physiological perspective, this pretended dataset is valid and represents the escape panic because in the real escape panic the value of HR and movement will be almost the same or mostly much higher (due to the release of more adrenaline and nor-adrenaline hormones). (2) To perform normal daily activities (without heavy physical activities such as playing sport). (3) We found that climbing stairs or a high hill will potentially increasing the heart rate while it is not a panic case. So we collected such data to enable the machine learning model to learn the correlation between the heart rate and the movement of climbing stairs as a normal activity. (4) On the other hand, we manually synthesized 112 abnormal activity samples that reflect the freezing action; when frightened people are freezing in their place and do not escape; but for sure their heart rate is pumping very fast. So we produced these synthetic datasets and we took in our consideration the following factors; First, we set the HR values to be

extraordinary from normal, at the same time do not exceed $(220 - \text{age of a person})$ [29]. Second, the respective movement data was very minimal because during the freezing action, people had limited motion.

Labelling the normal and abnormal cases was done via a handcrafted labelling process. The Full details of data segmentation and pre-processing are in the following sections.

D. Data Pre-processing

1) Feature Selection from Row Data

Synchronized raw data from the wristband sensor (HR-monitor and movement data) is merged into 1 data file per subject per session, available as CSV-files (.csv). Each of the data-files contains 5 features/columns, and 70-time steps/rows, the columns contain the following data:

- 0 Timestamp (s)
- 1 Age
- 2 Gender
- 3 Weight
- 4 Movement
- 5 Heart Rate (bpm)

2) Data Segmentation

To segment the data, we used a sliding window of fixed length; the length of the window is 70s, with 1-second interval. The number of instances (segments) obtained after using this sliding window configuration is 1054 data samples (532 are normal and 522 are abnormal using handcrafted labeling).

The dataset of this study is a time series which is a sequence of real-valued data points with timestamps. We denote a time series as $T = \{t_1, t_2, \dots, t_n\}$, where t_i is the HR and motion values at time stamp i and there are n timestamps for each time series ($n=70$ seconds). In our study, our assumption is that all training series have the same number of timestamps. We denote a labeled time series dataset as $D = \{(T_j, y)\}_{j=1}^N$ which contains N time series data ($N=1054$) and their associated labels. For each $j=1 \dots N$, T_j represents the time series and its label is y . Our classification problem is binary

where y is a binary value $y = \{0, 1\}$ the label 0 is an indication that it is normal (no panic) and the label 1 is an abnormal situation.

It is important to emphasize that the sequence length ($n = 70s$) was not arbitrarily chosen; according to our experiments, we came up with the following justifications: (1) We ran our experiment with ($n = 70s, 80s, 90s, 100s$). We realized that the accuracy was better with $n=70s$ and $n=80s$, but we have chosen 70s because our system's objective is to achieve swift detection and it is better to make it as small as possible. On the other hand, we excluded any values that are less than 70s, because it may be computationally expensive.

3) Normalization

Artificial neural networks are one of the machine learning models that need data to be normalized, so we performed min/max normalization approach.

E. Using Deep Learning Sequence Classification Models for Mass Panic Detection

Deep Learning is a subfield of machine learning. It uses a layered structure of artificial neural networks (ANN). In recent years deep learning has attracted tremendous research attention and proven outstanding performance in many applications compared to conventional machine learning algorithms [30], [31].

Recurrent Neural Networks (RNNs), such as Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU), have achieved a significant outstanding performance over traditional machine learning models. RNNs have the ability to capture long-term temporal dependencies and variable-length observations and use contextual information when mapping between input and output sequences. They are able to deal with and overcome the vanishing gradient problem. (LSTM) was introduced by Hochreiter and Schmidhuber (1997) [32]. While (GRU) was recently proposed by Cho et al. [33], it is like LSTM but with fewer parameters.

According to our data collection and experiments, it was very difficult to find an empirical panic threshold that works with every human, due to the physiological fact that, every human being has different heart rate beating zones, as heart rate varies from person to person. There are many interior and exterior factors that can cause heart rate fluctuation (e.g. emotional or physical exertion will speed up the pulse). This depends on several factors such as age, gender, weight, activity, and fitness [34]. So we decided to use a machine learning classification approach.

Machine learning models have attracted both academia and industry entities and has proven outstanding performance in several applications. They are able to learn, adapt, and perform good classification. Furthermore, the detection itself is considered as a sequence classification and labeling problem. In order to differentiate whether there is a panic or not, there is a need to analyze the data as a time series (learning the relationship between the increment of HR and the respective accelerated movement and trace the situation for a specific time. To this end, it was necessary to handle this detection using a sequence deep learning model like RNN.

In this work, supervised deep learning sequence classification models are used, LSTM and GRU (Type: many to one). To run experiments on RNNs models, we used drop-out with a value of 0.2. Number of epochs = 50 iterations. The internal architecture of LSTM and GRU is one layer consisting of 32 nodes and a time step of the sequences ($n=70$). Fig. 5. illustrates the used machine learning models.

F. Using Non-Sequential Machine Learning Models

For the sake of comparison, we also used the non-sequential machine learning models one-class SVM with default parameters, and Vanilla Neural Network (VNN). Our dataset is a time series being of equal length (time steps=70). The classification can be done using RNNs as well as VNN and SVM, so we compared the performance of sequential and non-sequential models. Note that the data representation in the non-sequential models is by inserting the data one after another so it is a one dimension vector of size 350 (5 features * 70 time steps).

G. Impelementaion Environmnet

We performed our experiments in Ubuntu environment using Keras [35] (on top of TensorFlow [36]) using Python (version 2.7). In this study, the models training and classification are performed on Intel (R) Core (TM) i7-3630QM CPU @2.40MGz and GPU NVIDIA Quadro K2000M With 8GB of shared GPU memory, and 16 GB RAM.

H. Results and Evaluation Metrics

To evaluate this study, we split the dataset ($N=1054$) into a training set (70%) which was used to train the model on k-fold cross validation $k=5$ and we left out 30% of the dataset for testing (unseen data). Accuracy metric represents the percentage of correct classification. We got accuracy of (97.48) % using LSTM, (95.58)% using GRU, (94.32)% using SVM and (94.01)% using VNN. Also, we calculated the False Positive Rate (FPR) which is the percentage of normal activities that are detected falsely as abnormal activities. Also we calculated the False Negative Rate FNR, as the percentage of abnormal activities that are classified falsely as normal activities (see Table I).

TABLE I. RESULTS OF TEST EXPERIMENTS (30% UNSEEN DATA)

Model		False Positive Rate (FPR)	False Negative Rate (FNR)	Accuracy (30% Unseen Data)
Sequential	LSTM	1.83 %	3.27 %	97.48 %
	GRU	6.71 %	1.96 %	95.58%
Non-Sequential	SVM	4.27%	7.19 %	94.32%
	VNN	9.76 %	1.96 %	94.01%

A confusion matrix is summarizing the performance of the classification model on a set of test data, it gives a better idea of what types of errors the classifier is making. It provides a good visualization of the performance of a classifier. Table II-V are summarizing the performance of LSTM, GRU, SVM, and VNN respectively.

VI. DISCUSSION

Regarding our experiments, we found that using sequence labeling methods such as LSTM and GRU can capture the temporal and spatial relationship between the HR and movement readings, compared to non-sequential models like SVM and VNN. Our results show that LSTM performed the best, followed by GRU, while SVM and VNN achieved less accuracy. Note that a small false positive rate (FPR) is acceptable because there is still data analysis phase II that can handle the false positives and differentiate between the real panic and normal HR and motion increment. Furthermore small false negative rate (FNR) is also acceptable because this false detection happens at an individual level, not for all attendees.

It is important to emphasize that our proposed system suits any mass gathering that does not involve extreme movement or excitement. To apply our system to such excited gatherings, data collection should be involved in a way that emulates the normal and abnormal activities for that particular event.

Furthermore, we discuss the limitations of our proposed system in the followings:

1) *Cost of IoT Wearable Wristband:* While the cost of wearable devices is declining significantly, still there is a significant cost in buying a huge number of sensors. (Although distributed over a large number of people as part of the event's cost). Perhaps when wristbands become standardized and become a common gear for everyone like cellular phones, this limitation will be lifted and people will participate in such mobile crowd sensing using their own wristbands.

2) *Deployment in Portable Devices with Limited Resources:* Power consumption and battery drain are very critical challenges because embedded wristbands have low computational/memory/energy resources. At the present, there are substantial challenges in deploying and performing inference of deep learning models in wearables. However, recently Google has announced TensorFlow Lite and TensorFlow Mobile [37]. They are two lightweight solutions for deploying AI on mobile and embedded devices. Furthermore, these lightweight solutions are powered by small hardware chips such as NVIDIA's Jetson TX2 and USB-based Intel's Movidius. Although these powerful chips still cost significantly. In the recent future, their prices will fall off and they will be used at large-scale.

3) *User Acceptance and Willingness to Contribute:* Our proposed system requires the obligatory or voluntary participation of thousands or even millions of individuals. First, individuals need to be persuaded or obliged to buy and/or carry the suitable devices. For example, in Hajj the Saudi Ministry of Hajj has already begun considering a plan to make electronic bracelets mandatory for all pilgrims that contain their personal data, according to ARAB NEWS [38].

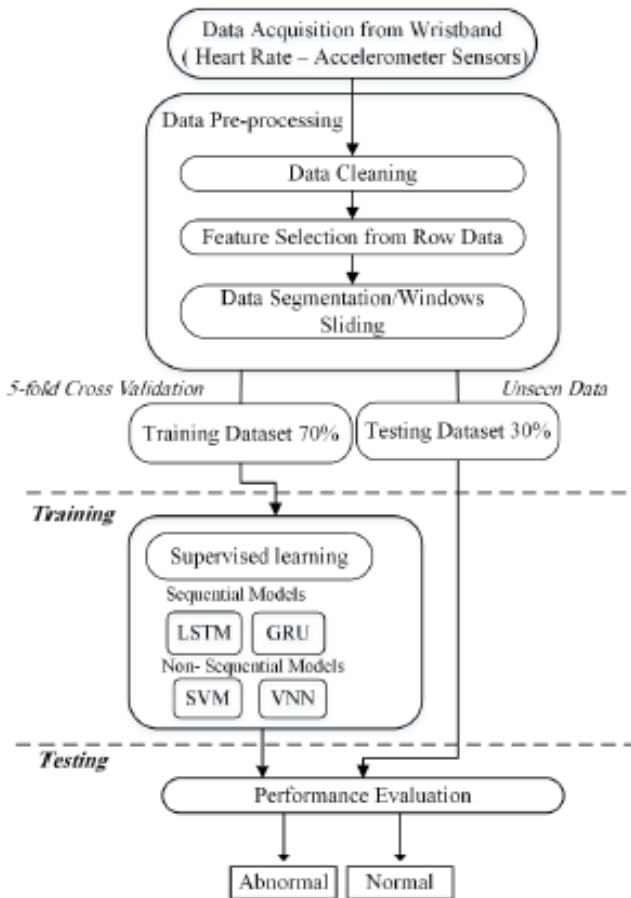


Fig. 5. Machine learning classification models.

TABLE. II. CONFUSION MATRIX OF TEST RESULTS (LSTM)

	Predicted	
	Abnormal	Normal
Actual		
Abnormal	148	5
Normal	3	161

TABLE. III. CONFUSION MATRIX OF TEST RESULTS (GRU)

	Predicted	
	Abnormal	Normal
Actual		
Abnormal	150	3
Normal	11	153

TABLE. IV. CONFUSION MATRIX OF TEST RESULTS (SVM)

	Predicted	
	Abnormal	Normal
Actual		
Abnormal	142	11
Normal	7	157

TABLE. V. CONFUSION MATRIX OF TEST RESULTS (VNN)

	Predicted	
	Abnormal	Normal
Actual		
Abnormal	150	3
Normal	16	148

VII. CONCLUSION AND FUTURE WORK

IoT and mobile crowd sensing applications help us to feel, see, and hear things that we never imagined before. This paper has proposed a new mobile crowdsourcing real-time detection method for abnormal crowd behaviour in mass gatherings. This system is based on advanced wireless and wearable sensors technologies and communication networks. It considers the humans themselves as the surveillance devices that exist everywhere. The proposed approach can detect mass panic swiftly in real-time manner by detecting the heart rate increase and abnormal motion.

In this study, we collected our own dataset from 89 subjects wearing an electronic wristband that has HR and motion sensors and we generated 1054 dataset samples. In this proposed approach, two-phases of data analysis were involved. Phase-I is a deep machine learning model that was used to analyze the sensors' collected readings of the wristband and detect if the person has indeed panicked so then sending alerting signals. While phase-II data analysis takes place in the monitoring server that receives alerting signals to conclude if it is a mass panic incident or a false positive case. The experimental results indicate that our developed deep learning sequential models LSTM and GRU got good accuracy of 97.48%, 95.58%, respectively. Compared to non-sequential models achieved only 94.32% for SVM and 94.01% for VNN.

In the future work, we will further develop data analysis phase-II module. Furthermore, using the same dataset, we will develop another detection algorithm that depends on anomaly detection approach; where we will train the RNN on normal data and predicting abnormality if there is a deviation of normality. Furthermore, we are planning to add another sensor, which measures skin conductance response (Electrodermal Activity EDA) that increases in sympathetic responses.

ACKNOWLEDGMENT

We would like to acknowledge participating subjects in the data collection phase for their willingness to contribute to this research and for the effort and time dedicated to data collection.

REFERENCES

- [1] L. Soomaroo and V. Murray, "Disasters at mass gatherings: lessons from history," *PLoS currents*, vol. 4, 2012.
- [2] B. Guo, Z. Yu, X. Zhou, and D. Zhang, "From participatory sensing to mobile crowd sensing," in *Pervasive Computing and Communications Workshops (PERCOM Workshops)*, 2014 IEEE International Conference on, 2014, pp. 593-598.
- [3] Egham. (2017, Accessed 29/11/ 2017). Gartner Says 8.4 Billion Connected "Things" Will Be in Use in 2017, Up 31 Percent From 2016. Available: <https://www.gartner.com/newsroom/id/3598917>
- [4] S. C. Mukhopadhyay, "Wearable sensors for human activity monitoring: A review," *IEEE sensors journal*, vol. 15, pp. 1321-1330, 2015.
- [5] A. Kumar, "Panic Detection in Human Crowds using Sparse Coding," 2012.
- [6] D. Helbing, I. Farkas, and T. Vicsek, "Simulating dynamical features of escape panic," *Nature*, vol. 407, pp. 487-490, 2000.
- [7] H. S. Bracha, "Freeze, flight, fight, fright, faint: Adaptationist perspectives on the acute stress response spectrum," *CNS spectrums*, vol. 9, pp. 679-685, 2004.
- [8] N. B. Schmidt, J. A. Richey, M. J. Zvolensky, and J. K. Maner, "Exploring human freeze responses to a threat stressor," *Journal of behavior therapy and experimental psychiatry*, vol. 39, pp. 292-304, 2008.
- [9] J. Leach, "Why people 'freeze' in an emergency: temporal and cognitive constraints on survival responses," *Aviation, space, and environmental medicine*, vol. 75, pp. 539-542, 2004.
- [10] A. R. Mawson, "Understanding mass panic and other collective responses to threat and disaster," *Psychiatry: Interpersonal and biological processes*, vol. 68, pp. 95-113, 2005.
- [11] A. Haag, S. Goronzy, P. Schaich, and J. Williams, "Emotion recognition using bio-sensors: First steps towards an automatic system," in *Tutorial and research workshop on affective dialogue systems*, Springer, 2004, pp. 36-48.
- [12] K. Kozłowska, P. Walker, L. McLean, and P. Carrive, "Fear and the defense cascade: clinical implications and management," *Harvard review of psychiatry*, vol. 23, p. 263, 2015.
- [13] T. Li, H. Chang, M. Wang, B. Ni, R. Hong, and S. Yan, "Crowded scene analysis: A survey," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 25, pp. 367-386, 2015.
- [14] B. Zhan, D. N. Monekosso, P. Remagnino, S. A. Velastin, and L.-Q. Xu, "Crowd analysis: a survey," *Machine Vision and Applications*, vol. 19, pp. 345-357, 2008.
- [15] M. Crocco, M. Cristani, A. Trucco, and V. Murino, "Audio surveillance: a systematic review," *ACM Computing Surveys (CSUR)*, vol. 48, p. 52, 2016.
- [16] R. Barnes and B. Rosen, "911 for the 21st century," *IEEE Spectrum*, vol. 51, pp. 58-64, 2014.
- [17] L. K. Moore, *Emergency communications: The future of 911*: Congressional Research Service, 2009.
- [18] T. Cheng and T. Wicks, "Event detection using Twitter: a spatio-temporal approach," *PloS one*, vol. 9, p. e97807, 2014.
- [19] T. Simon, A. Goldberg, and B. Adini, "Socializing in emergencies—A review of the use of social media in emergency situations," *International Journal of Information Management*, vol. 35, pp. 609-619, 2015.
- [20] A. A. Press, "CSIRO software monitors Twitter to guide emergency services to fires," in *theguardian*, ed, 2013.
- [21] S. O. Goldman, R. E. Krock, K. F. Rauscher, and J. P. Runyon, "Security system triggered by heart rate detection," ed: Google Patents, 2013.
- [22] D. W. McCleary, S. M. Rosato, J. O. Uchidiuno, W. Xiyang, and J. D. Weisz, "Detecting and notifying of various potential hazards," ed: Google Patents, 2018.
- [23] T. Wada, H. Higuchi, K. Komaki, H. Iwahashi, and K. Ohtsuki, "Disaster Detection Using SVDD Group Learning for Emergency Rescue Evacuation Support System," *Journal of Advanced Simulation in Science and Engineering*, vol. 3, pp. 79-96, 2016.
- [24] H. Iwahashi, H. Higuchi, K. Kogo, T. Kitamura, T. Wada, H. Okada, and K. Ohtsuki, "Extraordinary Judging Using Electrodermal Activity and Acceleration for Emergency Rescue Evacuation Support System," in *Parallel Processing Workshops (ICCPW)*, 2014 43rd International Conference on, 2014, pp. 355-360.
- [25] D. E. Long, *The Hajj Today: A Survey of the Contemporary Pilgrimage to Makkah*: SUNY Press, 1979.
- [26] D. Helbing, A. Johansson, and H. Z. Al-Abideen, "Dynamics of crowd disasters: An empirical study," *Physical review E*, vol. 75, p. 046109, 2007.
- [27] TOMTOM SPARK CARDIO WRISTBAND. Available: <https://www.tomtom.com>
- [28] T. Tamura, Y. Maeda, M. Sekine, and M. Yoshida, "Wearable photoplethysmographic sensors—past and present," *Electronics*, vol. 3, pp. 282-302, 2014.
- [29] H. Tanaka, K. D. Monahan, and D. R. Seals, "Age-predicted maximal heart rate revisited," *Journal of the American College of Cardiology*, vol. 37, pp. 153-156, 2001.
- [30] J. Schmidhuber, "Deep learning in neural networks: An overview," *Neural networks*, vol. 61, pp. 85-117, 2015.

- [31] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, p. 436, 2015.
- [32] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural computation*, vol. 9, pp. 1735-1780, 1997.
- [33] J. Chung, C. Gulcehre, K. Cho, and Y. Bengio, "Empirical evaluation of gated recurrent neural networks on sequence modeling," arXiv preprint arXiv:1412.3555, 2014.
- [34] Heart.org. (July 2015, Accessed: 24/5/2017, 11:30 pm). American Heart Association, All About Heart Rate (Pulse) Available: <http://www.heart.org/>
- [35] F. Chollet. (2015, Accessed:13/6/2016). Keras. Available: <https://keras.io/>
- [36] Google. (2015, Accessed: 11/5/2016). Available: <https://www.tensorflow.org/>
- [37] TensorFlow. (2018, Accessed:15/2/2018). TensorFlow Lite versus TensorFlow Mobile. Available: <https://www.tensorflow.org/mobile/>
- [38] ARABNEWS. (2016, Accessed: 21/8/2016). E-bracelets a 'must' for pilgrims this Haj. Available: <http://www.arabnews.com/node/947221/saudi-arabia>