Preprint submitted to the Springer Natural hazards journal. You can find the published version at:
https://link.springer.com/article/10.1007/s11069-024-06547-8
View-only version: https://rdcu.be/dDRPD

Harnessing trustable crowdsourcing power for flood disaster evaluation

Sajedeh Abbasi, Pervasive and Cloud Computing Lab, Faculty of Electrical and Computer Engineering University of Birjand, Birjand, Iran (Sajedeh_abbasi@birjand.ac.ir)

Hamed Vahdat-Nejad*, Pervasive and Cloud Computing Lab, Faculty of Electrical and Computer Engineering University of Birjand, Birjand, Iran (Vahdatnejad@birjand.ac.ir) (Corresponding author, ORCID id: 0000-0002-1478-2686)

Hossein Moradi, Birjand University of Technology (Hsmoradi@birjandut.ac.ir)

***Abstract*— This paper presents a proposed information and communication technology-based system that uses a crowdsourcing model to collect and provide accurate and up-to-date information about flooded areas. The system aims to assist relief organizations to act more efficiently following a flood disaster. The system collects data related to four informational requirements: people and animals, facilities for living, medical facilities, and shelters and roads. The proposed system includes a malicious user detection algorithm to prevent inaccurate information and keep the data current. The paper also introduces an information aggregation algorithm and a user reputation score algorithm to identify high-scoring users. The three proposed algorithms are assessed using simulation, which shows that they can accurately identify malicious users and rank non-malicious users. By providing up-to-date information from flooded areas, the system can help relief organizations respond more effectively to a flood disaster.**

***Keywords*— Disaster relief, Flood, Mobile Crowdsourcing/Crowdsensing, Information acquisition.**

I. INTRODUCTION

Every year, news is reported on the incidence of natural disasters in different geographical areas. In addition to causing human and financial losses, natural disasters create numerous problems for disaster victims, such as a lack of access to basic necessities, and they make living conditions difficult for survivors (See, 2019). Floods are the most prevalent type of natural disaster with a life-threatening nature. Urban sprawl along rivers and climate change are causing devastating floods (Chen et al., 2017; Ramesh et al., 2022). Flood disasters account for more than half of the fatalities and one-third of the economic damage resulting from natural disasters (Neumayer, Plümper, & Barthel, 2014). To effectively manage flood disasters, it is necessary to consider prevention, preparedness, response, and recovery (Abrahams, 2001; Cronstedt, 2002). The purpose of this research is to address information acquisition for the response phase.

It is crucial for the relief teams to have accurate and timely information about the needs of the victims to plan response and rescue operations effectively (Caballero-Anthony, Cook, & Chen, 2021; Suri et al., 2018). Having this information can help improve the quality of relief efforts and potentially expedite the process. Eyewitnesses play a crucial role in providing real-time information to rescue forces. In this regard, smartphones can play a vital role, as they are convenient and effective for acquiring and transmitting these observations quickly and efficiently. With smartphones' capability to capture photos, videos, and audio recordings, eyewitnesses can document the situation and relay important details to the rescue forces.

With the widespread availability of smartphones, individuals can now contribute to data collection efforts through crowdsourcing. Crowdsourcing involves the participation of a group of people in dynamically collecting and sharing information (Estellés-Arolas & González-Ladrón-de-Guevara, 2012; Vahdat-Nejad, Asani, Mahmoodian, & Mohseni, 2019). This information can provide insights into the prevailing conditions of the disaster area and aid in understanding environmental patterns and changes. Previously, crowdsourcing has been successfully applied in various fields, such as medicine (Golumbic et al., 2023; Tucker, Day, Tang, & Bayus, 2019), engineering (Mao, Capra, Harman, & Jia, 2017), tourism (Shi, Zhao, & Chen, 2017), trade (Kohler, 2015), and disaster response (Hultquist & Cervone, 2020; Sermet, Villanueva, Sit, & Demir, 2020; Suri et al., 2018; Vahdat-Nejad et al., 2019; Vahdat-Nejad, Bahadori, & Abiri, 2021). In the context of relief forces, crowdsourcing can provide a cost-effective and efficient way to acquire information from disaster environments (Ludwig, Siebigteroth, & Pipek, 2014). Flood victims, who have firsthand experience of the conditions in the region, can contribute valuable information to the data collection process.

During a flood, the probability of losing all mobile antennas is generally low. Mobile network operators often have backup systems and contingency plans in place to ensure that communication infrastructure remains operational during such events. However, it is important to note that severe flooding or other natural disasters can still cause disruptions in mobile network coverage. However, the adoption of satellite internet technology can indeed provide a more reliable and resilient means of communication during emergencies, including floods (Kodheli et al., 2020; Zuo et al., 2023). As a result, even if terrestrial Internet connectivity is disrupted, the system can continue to function effectively, as demonstrated by several studies on collecting flood information (Eckhardt, Leiras, & Thomé, 2022; Frigerio et al., 2018; Sahay, Kumar, Pongpaichet, & Jain, 2017). Furthermore, crowdsourcing methods enable people in the disaster zone to send their information to relief teams during and after the disaster. It is not necessary for everyone involved in the disaster to use the application; even a small fraction of users can provide enough valuable and reliable data rapidly.

Crowdsourcing has limitedly been used to collect necessary information about floods. In this regard, MAppERS allows users to upload photographs taken from the water level and send information about the flood and their requirements to rescue forces (Frigerio et al., 2018). In another crowdsourcing system (Sahay et al., 2017), efforts have been made to connect people and resources in emergency situations. To this end, they used multimedia crowdsourced information, while the main goal is to connect people and required resources, which is time-consuming. In crowdsourcing systems, it is crucial that the information transmitted by users is correct and up-to-date (Gao, Barbier, & Goolsby, 2011; Wu et al., 2023). However, despite the best efforts of relief forces, the use of incorrect information sent by users can mislead the rescue forces, causing delays in aid and potentially costing lives (Goolsby, 2010; Vadavalli & Subhashini, 2023). The research gap identified in the investigated research projects is the lack of emphasis on the comprehensiveness and accuracy of the collected flood data. The focus is primarily on collecting a limited type of information without verifying data accuracy. However, it is crucial to ensure the accuracy of the crowdsourced data because inaccurate information can lead to erroneous conclusions and misguide rescue forces.

This paper uses crowdsourcing to propose a comprehensive data (including text, audio, images, and multiple-choice questions) collection system for flood disasters. It is based on our previous categorization (Abbasi, Vahdat-Nejad, & Hajiabadi, 2022) of the essential information requirements of the area, which are people & animals, living facilities, medical facilities, and shelters & roads. This research also addresses the issue of malicious users of crowdsourcing systems by detecting and removing their data. It utilizes statistical methods to expedite the detection of malicious users, thereby increasing confidence in the compiled data and facilitating relief operations. To augment users' credibility, we incorporate training and hardware parameters alongside behavioral parameters to obtain users' reputation scores.

The contributions of this research can be summarized in three main aspects:

- **New reference model for crowdsourcing in the flood domain:** The research presents a novel reference model specifically designed for crowdsourcing in the flood domain. Although limited models (e.g., a model that only gathers multimedia (Sahay et al., 2017)) were proposed earlier, our reference model takes into account the unique challenges and requirements of managing crowdsourced data related to floods. Besides, the proposed reference model contains the required modules for guaranteeing the correctness of information. By proposing this new reference model, the research aims to improve the effectiveness and efficiency of crowdsourcing efforts in flood management.
- **Malicious user detection algorithm:** The research introduces a sophisticated algorithm for detecting malicious users in the crowdsourcing system. This algorithm not only identifies users who engage in malicious activities, such as providing false or misleading information but also grants the system the ability to delete data and remove the malicious user from the platform. This capability enhances the reliability and trustworthiness of the crowdsourced data.
- **User reputation scoring algorithm:** The research proposes a user reputation scoring algorithm that takes into account various sub-parameters of a user's past behavior and their training and hardware characteristics. This approach helps to identify the most useful users within the crowdsourcing system.

Two scenarios implemented in MATLAB were developed to evaluate the proposed algorithms. The evaluation results validate the correct operation of the proposed malicious user detection algorithm and demonstrate the appropriate ranking of users by the proposed user reputation score algorithm. Consequently, these algorithms serve as effective tools for preventing the collection of incorrect information.

This article proceeds as follows. In Section II, we categorize and review related studies. Section III introduces the proposed crowdsourcing system. Section IV describes the malicious user detection algorithm, while Section V discusses the user reputation score algorithm and the information aggregation algorithm. The evaluation of the proposed algorithms is presented in Section VI. Finally, Section VII provides concluding remarks.

II. RELATED WORK

A variety of crowdsourcing and data collection projects have been presented in the past. We categorize the reviewed research into two categories: "flood information collection" and "user credibility in the crowdsourcing system." These systems have commonly relied on mobile sensors and the Internet for data collection and detection purposes. Previous user credibility research in the crowdsourcing system has been facilitated by the use of mathematical functions and the development of new components. The related articles are described in relation to these topics.

A. Collecting flood information

Utilizing user-submitted images from flooded areas is one way to detect flooding. In this regard, an image processing pipeline has been presented to detect floods by collecting and processing images (Witherow, Elbakary, Iftekharuddin, & Cetin, 2017). Similarly, the possibility of flooding was predicted by using a neural network that processes submitted images and mobile phone sensor data (Wang, Mao, Wang, Rae, & Shaw, 2018). Furthermore, the 3-cone intersection method was employed to develop a monitoring and forecasting system for flooding (Sermet et al., 2020). Another study utilized user-submitted data and mobile sensor data to measure water levels, aiding in system forecasting and flood risk map creation (Burkard, Fuchs-Kittowski, & de Bhroithe, 2017). FLOODIS (Rossi et al., 2015) is a cloud-based Service-Oriented Architecture in which users record flood information. It is then integrated with flood forecasting data from the European Flood Awareness System (EFAS), providing a Decision Support System (DSS) for government management.

Additionally, the Mapster crowdsourcing system (Liu et al., 2011), which collects and processes tweets from a geographic region along with radar data, has been implemented. A Bing-based Spatiotemporal visualization tool is used to map information from both data collection methods and predict floods by analyzing previous meteorological events. In this context, a method for creating a road damage map has been proposed (Schnebele, Cervone, & Waters, 2014) by combining aerial photographs of the flooded area, YouTube geographic videos, and tweets published during the flood period.

Four data sources have been analyzed for collecting information on floods, including sensor networks, satellite remote sensing imagery, hydrodynamic modeling, and crowdsourcing (Hultquist & Cervone, 2020). Upon completion of the research, it was determined that the crowdsourcing model does not have the problems associated with the aforementioned methods. Using the Internet of Things (IoT) is another method of data collection. In this regard, the authors (Suri et al., 2018) utilize IoT devices to collect data on the disaster zone. Additionally, the Flooded Streets system (Naik, 2016) is designed to respond to flood disasters. It uses OpenStreetMap data to render area map information hosted on GitHub. Users can update the map information by clicking on flooded streets. The CrowdMonitor system (Ludwig, Siebigteroth, & Pipek, 2015) also employs a web platform and a mobile application to collect disaster response data, allowing users to contribute flood data through these methods. On the other hand, a multimedia rescue system (Sahay et al., 2017) has been proposed to collect flood-related data, where users enter information such as flood level, requirements, and available resources. They can then view the map details of flooded and safe areas.

Finally, MAppERS is an application designed to collect information on flooding (Frigerio et al., 2018). When a disaster occurs, users upload photos of the water level and their requirements. Trained users help collect data by locating the geographical location of the hazard on a map and conducting field surveys. The application is linked to a web dashboard that provides relevant forces with access to flood information.

B. User credibility in crowdsourcing

Another significant challenge in crowdsourcing systems is assessing the credibility of users (Yu, Shen, Miao, & An, 2012). It focuses on building trust in crowdsourcing systems by studying trust interactions between Service Consumers (SCs) and Service Providers (SPs). Authors (Victorino, Estuar, & Lagmay, 2016) utilized web-based and mobile applications to collect flood data, using meteorological station data as a reference for verifying the location and timing of the disaster. Nearest neighbor search algorithms and neighbor searching with a fixed radius were employed to validate the user-submitted data. Additionally, a reputation management system was proposed (Kaleem, Majeed, Khan, Afzal, & Bashir, 2015) that calculates the user's credit score using a sum and mean model. Furthermore, a reputation management model for crowdsourcing systems was developed (Allahbakhsh, Ignjatovic, Benatallah, Bertino, & Foo, 2012), which takes into account trust and fairness as factors for determining user reputation.

III. PROPOSED CROWDSOURCING SYSTEM

Acquiring the requirements of flood victims in the first hours of a disaster is complex and time-consuming. The crowdsourcing model helps gather the most important requirements of flood victims in the shortest time and at the lowest cost (Bahadori, Vahdat-Nejad, & Moradi, 2022). In this paper, we propose a data collection system from the flooded area using a crowdsourcing model. To achieve this, we first identify the requirements of flood victims in four main categories: people & animals, living facilities, medical facilities, and shelters & roads (Abbasi et al., 2022).

A. The requirements of flood victims

Relief teams should be aware of the requirements of flood victims before planning and initiating an operation so they can provide the best services to those affected. Initially, we reviewed the requirements of flood victims from available resources (Fienen & Lowry, 2012; Suri et al., 2018; Zhao et al., 2016), including the Red Cross and Red Crescent sites. To increase the accuracy of identifying these requirements, we conducted semi-structured interviews with experts from the Iranian Red Crescent Society (Abbasi et al., 2022). The interview has five main questions answered by the CEO of the Red Crescent Society and ten rescuers who have at least 15 years of work experience. Based on this research, the information required by relief forces is divided into four categories: People & animals, living facilities, medical facilities, and shelters & roads (Abbasi et al., 2022).

- People & animals: As a result of a flood, some people lose their lives, while many become injured or missed. Rescuers require accurate information about the deceased, injured, and missing individuals. Rescue teams must also be aware of the remaining population in different parts of the area in order to allocate resources effectively. Additionally, there may be elderly or disabled individuals, as well as livestock or pets in need of assistance. Therefore, rescue workers require information about the elderly, disabled individuals, livestock, and pets.
- Living facilities: In flooded areas, water often enters houses, creating emergency situations for people. Access to food, baby formula, drinking water, tents, blankets, warm clothing, diapers, and toiletries are among the most critical requirements of flood victims.
- Medical facilities: The injured individuals require access to medical facilities and doctors. Some people affected by floods may have pre-existing medical conditions and require specific medications that are not readily available due to the flood. Therefore, rescue forces must have information about the medicines required by the people and the injured.
- Shelters & roads: One of the most important requirements of flood victims is shelter. Homes in flooded areas are typically destroyed or uninhabitable due to the intensity of the flood. Relief teams require information on uninhabitable homes to provide a safe place with adequate facilities for flood-affected families. Rescuers also require information on destroyed and mudded roads, streets, channels, tunnels, bridges, etc.

B. The reference model

Figure 1 illustrates the reference model of the proposed crowdsourcing system. The reference model divides the system functions into its elements and shows the data flow between the elements (Fettke & Loos, 2003). Figure 1 illustrates the reference model layers, data streams, relationships, and elements' functions. The proposed Crowdsourcing system is divided into four main layers, including data collection, data preprocessing, data processing, and data aggregation and sharing.

The data collection layer acquires the requirements of the users in the flooded environment with the help of a pre-designed questionnaire. In the preprocessing layer, two algorithms, "flooded area zoning" and "malicious user detection," are implemented. The mentioned algorithms are described in the next section. In the data processing layer, the "user reputation scoring" and "information aggregation" algorithms are implemented, which are described in the fifth section. Finally, in the distribution layer, aggregated information is shared, and a list of requirements and medications, along with user reputation scores, is provided to relief teams and related organizations.

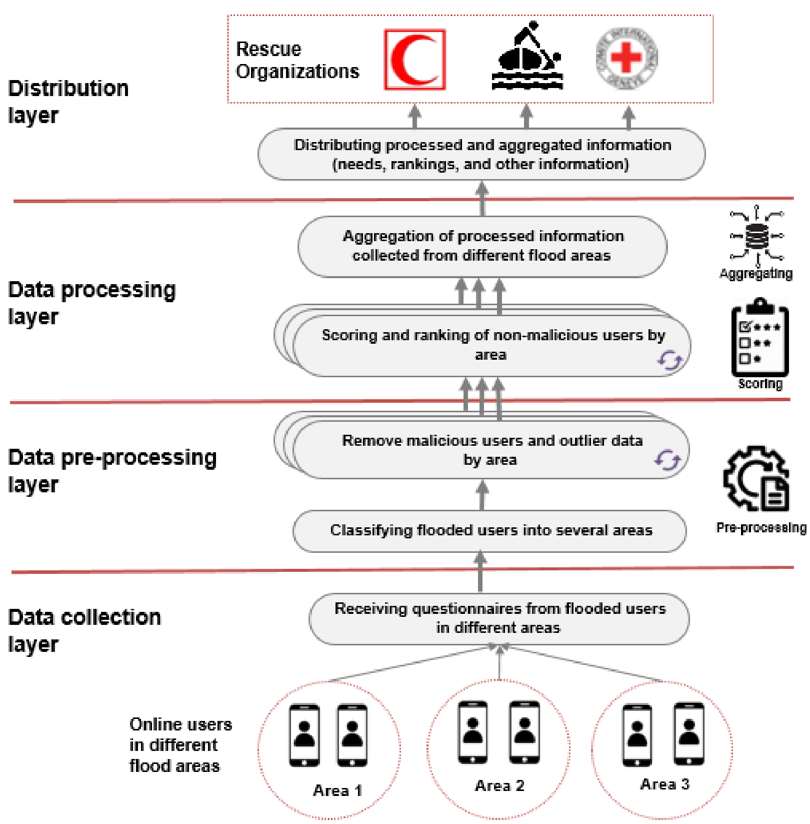

Figure 1: Reference model of the proposed Crowdsourcing system for the flood environment

A questionnaire with 15 multiple-choice items and two descriptive questions is designed for the data collection layer. By responding to the application's questions and sending them to relief teams, users help these teams obtain the information they require and expedite relief operations. However, analyzing descriptive questions is more time-consuming. As depicted in Figure 2, users enter their name, level of education, completed relief courses, and prior experience with relief operations after logging into the application.

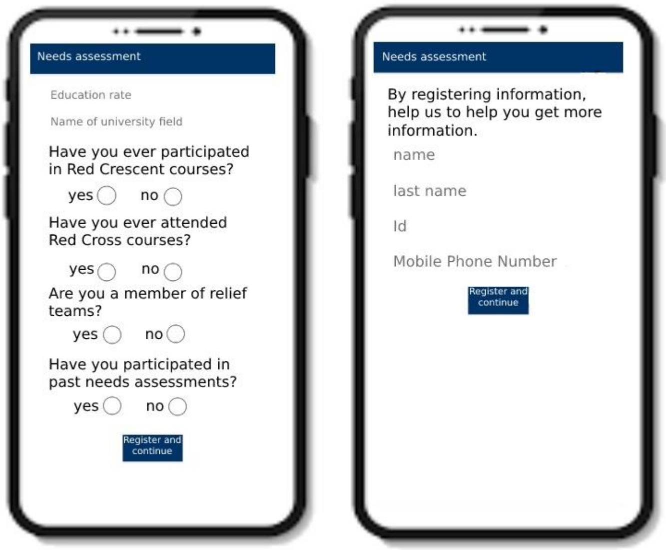

Figure 2: Receiving user information by the application

Figure 3 shows three typical pages of the proposed application. As it is evident, users can enrich the answers to each question by uploading audio, photo, or video files. They may also decide to skip any question. Users can type and send statements in response to descriptive questions. One of these questions (figure 3 (b)) asks a list of drugs required by users. In the final question (Figure 3(c)), users are free to provide any flood-related explanation.

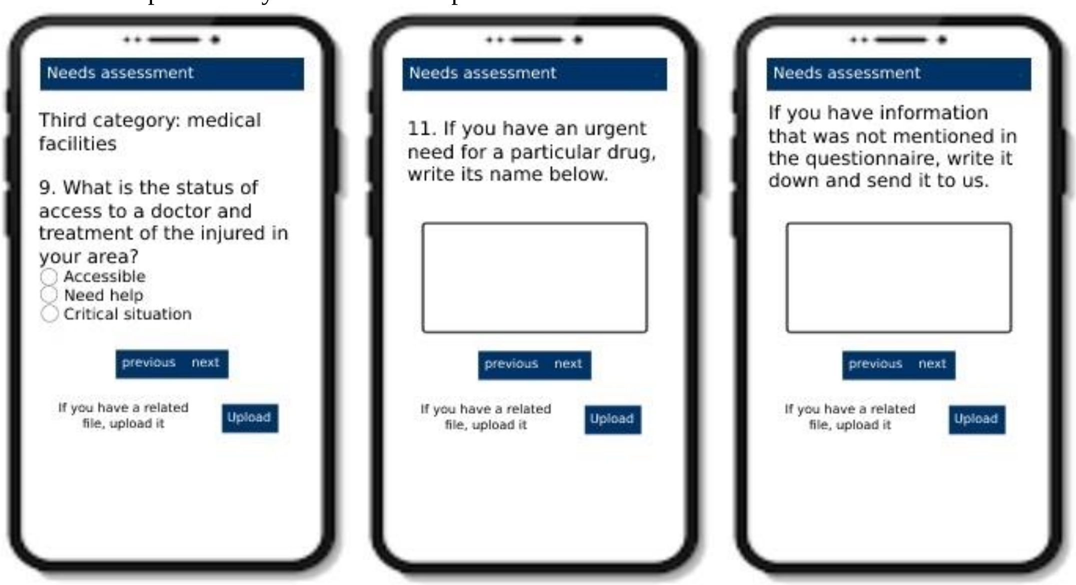

Figure 3: Typical screenshots of the proposed application. (a)A multiple-choice question. (b) A descriptive question. (c) Users enter additional information. (Abbasi et al., 2022)

IV. MALICIOUS USER DETECTION

In some instances, the crowdsourcing pattern is exposed to incorrect information, rendering the system unreliable. Indeed, one of the most significant challenges of a crowdsourcing system is the detection of malicious users (Yang, He, & Shi, 2016), as the transmission of incorrect information by malicious users can mislead rescue workers (Ludwig et al., 2014). To address this issue, we propose a two-part algorithm for malicious user detection, which includes data outlier detection and malicious user detection. The algorithm is a modification of our previous algorithm (Abbasi et al., 2022). In the proposed algorithm, outlier data is identified using statistical methods, and then malicious users are identified and removed from the system

To precisely identify user requirements, the disaster zone is partitioned into R sections (regions), similar to a chessboard, and the algorithm is executed separately for each section. Initially, each user response is digitized. For example, in the case of multiple-choice questions, the first option is assigned number one, the second is assigned number two, and so on. The same approach is used for three-choice and five-option questions. Afterward, the mean and standard deviation of the values of each question in each region are computed for each period (i.e., one hour), as follows:

$$\overline{x} = \left(\frac{1}{N}\right)\sum_{i=1}^{N} x_i \tag{I}$$

$$\delta = \sqrt{\left(\frac{1}{N}\right)\sum_{i=1}^{N}(x_i - \overline{x})^2} \tag{II}$$

Where N represents the number of data elements entered for that question, x_i denotes the values received from users, $\overline{x}$ is the average of the data received for each question, and δ is the standard deviation of the data pertaining to that question (Sievers, 2015). After calculating the mean and standard deviation of the received data at the end of a period, the interval [$\overline{x}$-2δ, $\overline{x}$+2δ] is considered valid. Therefore, if user data falls within this range, the user's response to the question will be considered valid (useful); otherwise, it will be deemed an outlier answer. The number of data outliers is then calculated for each user.

To prevent a malicious user from dominating the average value through incorrect contributions, we only factor one value per question from each user into the calculation of the average and deviations. If a user participates multiple times during a given period, the average of the data submitted by the user for each query is calculated and considered as the amount submitted by the user for that period. Equations (3) then detects outliers and malicious users:

$$A_i = \frac{Number\ of\ outlier\ data\ elements\ for\ user\ i}{Total\ sent\ data\ elements\ for\ user\ i} \tag{III}$$

$$\text{Participation type} \begin{cases} A_i <= 0.3 & \text{Non-malicious user} \\ A_i > 0.3 & \text{Malicious user} \end{cases}$$

When a user provides a certain rate of data outliers, it is possible to determine whether their participation was malicious or not by using Equation 3. The variable A_i represents the percentage of data outliers supplied by user i. Based on the evaluation performed and the results obtained, we determined that a malicious user is one whose submitted data contains at least 30% outliers. This ratio is obtained by using a 200-times testing and validation routine. This routine is explained in section VI. If a malicious user is detected, all information received from them will be removed from the system, and the user will no longer be considered for subsequent contributions. By eliminating malicious users and the data they submit, the information reaching the rescue forces becomes more reliable.

The malicious user detection algorithm is executed hourly at regular intervals. However, if fewer than five users participate in the system during this period, it will not be possible to correctly identify data outliers and malicious users. Therefore, the algorithm for detecting malicious users will not act. As such, the minimum number of statistical samples required for the malicious user detection algorithm to calculate the mean and standard deviation is 5.

V. User Reputation Scoring and Information Aggregation Algorithm

When it comes to gathering data for rescue workers, some users in mobile crowdsourcing systems are more helpful than others. However, due to the diverse user population in crowdsourcing systems, it is crucial to determine the credibility of users (Kaleem et al., 2015). In this regard, it is suggested to rank non-malicious users in order to identify the most useful ones. This ranking allows for the identification of users with higher ranks who can be contacted if additional information is needed. The user reputation score algorithm is implemented to rank non-malicious users, taking into account their performance, experience, and the hardware they use to transmit data. The parameters for user reputation scoring include "user's past behavior" and "training and hardware," each encompassing the aforementioned factors.

A. **User reputation scoring**

The key parameters and sub-parameters required for extracting the "user reputation score" are illustrated in Figure 4. The proposed user reputation scoring algorithm consists of two parameters: "user's past behavior" and "training and hardware." The parameter "user's past behavior" is considered one of the most crucial factors for ranking (Noorian & Ulieru, 2010; Yu, Shen, Miao, Leung, & Niyato, 2010). To quantitatively evaluate this parameter, criteria for comprehensiveness and usefulness of user participation have been proposed. Comprehensiveness assesses the extent to which the user responded to questions, while the usefulness shows the impact of the user in each region. The concepts of comprehensiveness and usefulness are elaborated later in this section.

The "training and hardware" parameter comprises three sub-parameters: training, Internet speed, and the quality of the user's mobile camera. The training sub-parameter evaluates the user's education and experience in critical situations. Since network coverage varies across regions, the Internet speed sub-parameter assigns an appropriate score based on the analysis of network coverage. Additionally, user scoring considers the variation in the quality of mobile phone cameras among users, recognizing that photos and videos are helpful in determining the requirements of flood victims. Continuing, we elaborate on computing all parameters and sub-parameters of the user reputation score. At the top level, the user's reputation score is calculated by adding the "past user behavior" and "training and hardware" parameters, as follows:

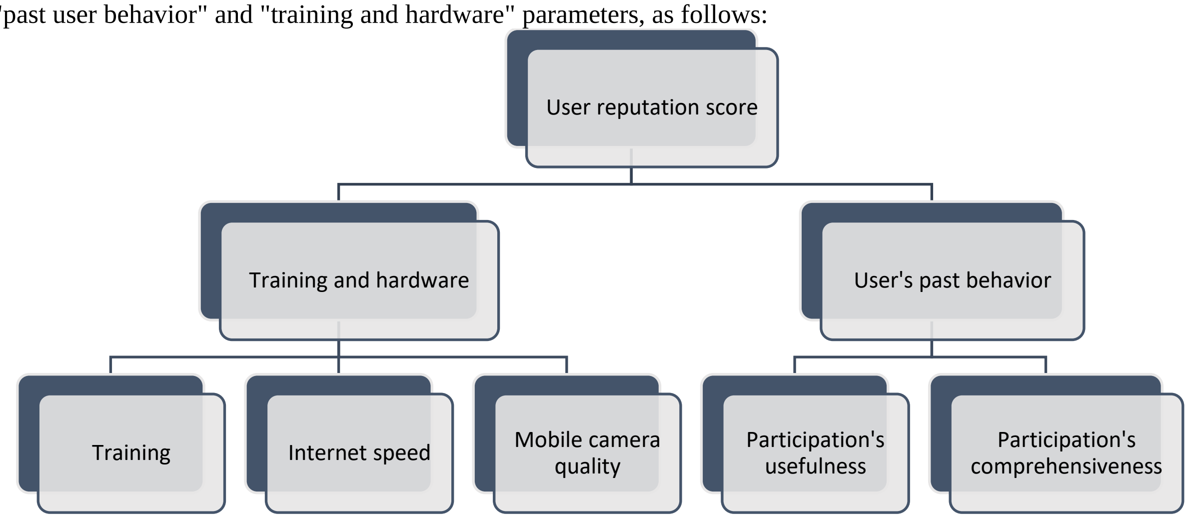

Figure 4: Factors determining the user reputation score for a flooded environment

$$User\ score\ =\ Past\ user\ behavior\ +\ training\ and\ hardware \qquad \text{(IV)}$$

User's past behavior: The value of the user's past behavior parameter is calculated by adding the values of the two sub-parameters, including comprehensiveness and usefulness of participation:

$$Past\ user\ behavior = Comprehensiveness\ of\ participation\ +\ Usefulness\ of\ participation \qquad \text{(V)}$$

Comprehensiveness of participation: When a user enters the application, they are presented with 15 multiple-choice questions related to information requirements. The user can choose to answer any of the questions or skip a question. A user who responds to more questions is considered more valuable to the rescue team than one who responds to fewer questions and thus should receive a higher score. Due to the zoned nature of the flooded area, the score for each user in each zone is calculated separately and ranges from 0 to 1. According to the following relationship, half of the scores of this sub-parameter belong to the participation percentage in the questions, and the other half is determined by the percentage of participation in the file uploading:

$$The\ comprehensiveness\ of\ user\ participation\ =\ \left(\frac{\sum_{i=1}^{k} Ni}{15K}\right) * 0.5 + \left(\frac{\sum_{i=1}^{k} Mi}{15K}\right) * 0.5 \qquad \text{(VI)}$$

Where N represents the number of responses to items in the user's i_{th} participation, M denotes the number of files (audio, photo, and video) uploaded in the i_{th} participation, and k indicates the number of user's participations in the last period. Reminding that 15 is the total number of questions, participation comprehensiveness simply calculates the response rate of the user. When all the results are summed up, a single number represents the extent of users' involvement, which falls within the interval [0, 1]. It should be noted that uploading numerous files might overload the server; however, cloud computing provides a huge capacity for computation and storage.

The usefulness of participation: As the flood zone is divided into distinct regions, there may be variations in participation levels across different zones. Consequently, there may be minimal data available from regions with lower participation. Therefore, users who provide information from these regions are considered more valuable to relief teams. Equations 7 and 8 are proposed to calculate users' usefulness in different regions. The parameter's minimum score is 0, while the maximum score can exceed 1.

$$Usefulness\ rate\ (i.r) = \frac{Number\ of\ unmalicious\ participations\ of\ user\ i\ in\ region\ r}{Total\ unmalicious\ participations\ of\ users\ in\ region\ r} \quad \text{(VII)}$$

$$Usefulness\ rate\ i = \sum_{r=1}^{R} usefulness\ rate(i.r) \quad \text{(VIII)}$$

The variable r represents a region, and R indicates the total number of regions. In Equation 7, the number of non-malicious participations of user i in the r^{th} region within the last period is divided by the total number of non-malicious user participation in the r region (over the last period). It determines the proportion of non-malicious participation in the r region attributed to user i. Equation 8 adds the usefulness rate of user i across all regions. Note that a user's score in this sub-parameter can exceed one if they contribute significantly from different regions.

Training and hardware: Evaluating the user's level of education is a crucial factor in boosting their score. However, consideration must also be given to the user's hardware, as it enables high-speed internet utilization and the capture of high-quality images. Therefore, the training and hardware parameter includes sub-parameters for training, internet speed, and camera quality. The scores for each of these sub-parameters are calculated using Equation 9 to obtain the training and hardware parameter score.

$$Training\ and\ hardware = Mobile\ camera\ quality + internet\ speed + training \quad \text{(IX)}$$

Training: During a disaster, trained and experienced users respond more effectively compared to those without proper training and experience. Consequently, these individuals are more valuable to relief teams. Table 1 is recommended for scoring the training sub-parameter.

Table 1: How to score the training sub-parameter to determine the user's score

Training and experience	Score
Passing Red Crescent courses	0.2
Passing Red Cross courses	0.2
Have a relevant university degree	0.2
Collaborate with relief teams	0.2
Participate in past Crowdsourcing operations	0.2
None	0

According to Table 1, users who are subject to one of the educational items mentioned in the table will receive a score of 0.2. Users who have a Red Crescent-approved score, such as first aid and preparedness against risks, receive a score of 0.2. Those who have completed Red Cross courses (like the HELP training course for managing relief operations during humanitarian crises) are also considered. Users with a university degree related to relief, such as disasters and urgency, management engineering in natural disasters, or crisis management, as well as medical expertise, receive a score of 0.2. Additionally, users who are members of the relief team currently present in the flooded region or have experience with crowdsourcing operations receive a score of 0.2. If none of these parameters apply to the users, their score will be zero. If multiple parameters apply, their training sub-parameter score will be the sum of the attained scores. The training score falls within the range of [0, 1]. In fact, a user can have none of the items in the above table or can have all of them. Hence, the maximum score obtained by a user is 1.

Internet speed: Users employ various transmission methods based on their hardware and internet coverage when sending data. Users with faster internet speeds can send information more quickly, while slow internet speeds may hinder data transmission. Table 2 assigns a score to each internet technology, which is used to calculate the user's score. The scores for internet speed fall within the interval [0, 1]. Here, the lowest score is given to a user who sends data using 3G mobile technology and obtains a score of zero. Moreover, based on the information provided in the table, the maximum internet score obtained by a user is 1.

Table 2: How to score users according to the Internet technology used by them

Internet type	Score
Mobile Internet (3G)	0
Wi-Fi	0.5
Mobile Internet (4G)	0.5
Mobile Internet (5G)	1

Camera quality: The quality of the user's mobile camera is also taken into account when calculating their score. A higher-quality camera enables better extraction of information from submitted photos and videos. Therefore, users with superior camera quality should receive a higher score. In this context, a camera resolution of at least 20 megapixels is considered the optimal standard for mobile devices, serving as the measurement unit for other devices as well. Equation 10 is proposed to score users based on their mobile phone camera quality. The appropriate score range for this sub-parameter is [0, 1].

$$\text{Camera resolution score} = \begin{cases} \text{If (Camera resolution} < 20) & \text{(Camera resolution/20)} \\ \text{Else} & 1 \end{cases} \quad \text{(X)}$$

Since the resolution of the user's camera is measured in megapixels, it is divided by 20. The maximum user score of the camera quality sub-parameter is 1. When the user's mobile camera has a resolution greater than 20 megapixels, the camera resolution score is set to 1.

B. Information aggregation

In each period, once the malicious user recognition algorithm and user scoring are executed, the data from non-malicious users is aggregated and sent to the relevant stakeholders. This algorithm compiles responses to multiple-choice questions and transmits them to relief teams. Additionally, descriptive answers are accumulated per question and region. The relief teams utilize this information to aid in the response to the flood. The following equations are employed to integrate answers to the multiple-choice questions:

$$Weight\ of\ user\ i\ = \frac{reputation\ score\ of\ user\ i}{The\ highest\ reputation\ score\ of\ users\ who\ have\ participated\ in\ the\ last\ period} \quad \text{(XI)}$$

$$New\ value\ for\ question\ j\ = \frac{\sum_{i=1}^{N} weight\ of\ user\ i * X(i.j)}{\sum_{i=1}^{N} weight\ of\ user\ i} \quad \text{(XII)}$$

The number of users who contributed within the last period is denoted by N, while the variable X(i,j) represents the answer choice made by user i in response to question j. Then, according to Equation 12, each question's updated value is obtained by calculating a weighted average among the responses of all users. If a user has participated more than once in the last period, their average response for each question (represented as X(i,j)) is considered.

Next, the average of the new value for each question, along with the previously obtained aggregated value for the same question, is considered the new aggregated value and will be sent to the relevant organizations:

$$Aggregated\ value\ = \frac{New\ value\ for\ question\ j + Old\ aggregated\ value\ for\ question\ j}{2} \quad \text{(XIII)}$$

Table 3 shows an example of aggregated results for a medical question (i.e., a descriptive question) focusing on the drugs required by flood victims. The relief forces require an estimation of the names and quantities of medications requested in each region. This table provides valuable information to rescuers regarding the types of medications required in each area.

Table 3: Medicines required by flood victims by region number and number of applicants

Name of drug	Region number	Number of applicants
Acetaminophen	5	3

VI. Evaluation

We developed two simulators using MATLAB to evaluate the effectiveness of the malicious user detection and the user reputation scoring algorithm. The evaluations of these algorithms are described in detail below.

A. Malicious user detection

In order to evaluate the malicious user detection algorithm, 14 types of simulated users are considered. These include a random user, a pattern-oriented user, an accurate user, and 11 users with responses following a normal distribution with variances ranging from 0.1 to 1.5. It is expected that random users and pattern-oriented users will be identified as malicious users. Additionally, a higher variance in user-submitted information is expected to increase the likelihood of being detected as a malicious user. Equation 3 is evaluated to assess the performance of the algorithm and the interval [$\overline{x}$-2δ, $\overline{x}$+2δ] is considered a valid range for the submitted data. The threshold of 30% for detecting malicious users in equation 3 was obtained by using a 200 times testing and validation routine. The 14 user types are described as follows:

User 1 (Random): Provides completely random responses.
User 2 (Pattern-oriented): Selects responses based on a specific pattern.
User 3 (Accurate): Answers all questions correctly.
User 4 to User 13: Respond according to a normal distribution, which has the correct answer as the mean. The variance varies from 0.1 to 1.

- **User 4** (Variance 0.1)
- **User 5** (Variance 0.2)
- **User 6** (Variance 0.3)
- **User 7** (Variance 0.4)
- **User 8** (Variance 0.5)
- **User 9** (Variance 0.6)
- **User 10** (Variance 0.7)
- **User 11** (Variance 0.8)
- **User 12** (Variance 0.9)
- **User 13** (Variance 1)

User 14 (Variance 1.5): Responds based on a normal distribution, which has the correct answer as the mean and variance of 1.5.
This evaluation scenario was repeated 200 times, and we computed the mean values, including the average number of data outliers, for each user type. The simulation results are depicted in Figure 5, which shows the percentage of outlier data for each user. A user is classified as malicious if their percentage of outlier data is greater than 30% (indicating at least five questions with outlier responses). Users with fewer outlier responses are considered non-malicious. According to this figure, the random user (user 1), the pattern-oriented user (user 2), and the user responding according to a normal distribution with a variance of 1.5 (user 14) have more than 30% outlier responses and are correctly identified as malicious users. Conversely, the accurate user (user 3) and users responding based on a normal distribution with variances ranging from 0.1 to 1 (users 4 to 13) have less than 30% outlier responses. Users who respond based on a normal distribution with high standard deviations tend to have a higher percentage of incorrect answers, as evident from the figure. Ultimately, malicious users are removed from the system, and their responses are omitted.
Reviewing the figure confirms that the evaluation results of the malicious user detection algorithm align with the anticipated predictions. As expected, the random and pattern-oriented users are correctly identified as malicious. Additionally, careless users whose responses were simulated by the normal distribution with high variance are correctly identified by the system as malicious.

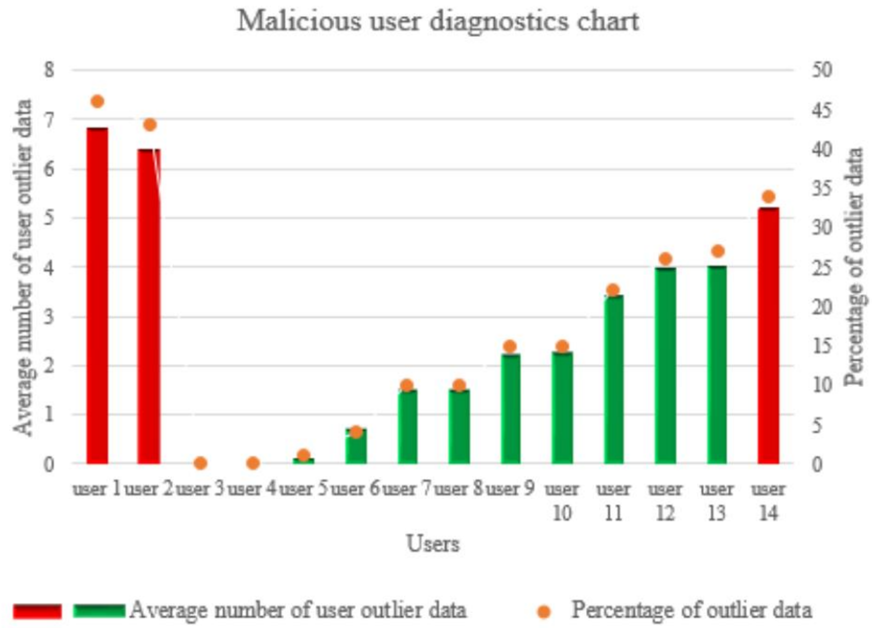

Figure 5: Evaluation of malicious user detection algorithm for 14 simulated user types

B. User reputation scoring

The reputation scoring algorithm incorporates factors such as the user's past behavior, training, and hardware. The user's past behavior is further divided into sub-parameters: participation's comprehensiveness and usefulness. The usefulness of participation metric suggests that users who contribute in areas with low overall participation are more valuable than those who contribute in areas with high participation. To assess the validity of this concept, we simulated three types of users as follows:

User 1: In Region 1, there are four participants, and User 1 is one of them.

User 2: In Region 2, there are two participants, and User 2 is one of them.

User 3: In Region 3, user 3 is the only participant.

To evaluate the impact of the participation usefulness parameter on the user reputation score, we require specific values to be assigned to the participation comprehensiveness and training and hardware sub-parameters. The participation comprehensiveness parameter is modeled using a uniform distribution function, assigning values ranging from 0.5 (indicating users who answered half of the questions) to 1 (representing users who answered all the questions). Additionally, the training and hardware parameter for each user is set to its maximum value of 1. According to Equation 8, the participation parameter values for users 1, 2, and 3 will be 0.25, 0.5, and 1, respectively.

The results of the simulation regarding user participation in the three regions are presented in Table 4. In regions with a higher number of contributions, the user receives a lower score. Conversely, in areas with fewer contributions, the user receives a higher score. This indicates that users who participate from regions where less information is available are deemed more useful to the relief teams and receive higher scores.

Table 4: Simulation result and the effect of user usefulness parameter on user score

User	Education	Comprehensiveness	Usefulness	**Reputation score**
#1	1	0.9786	0.25	**2.2286**
#2	1	0.7427	0.5	**2.2427**
#3	1	0.9001	1	**2.9001**

VII. CONCLUSION

In this paper, we proposed a reference model of the proposed crowdsourcing system for information gathering from flood disasters. The proposed architecture consists of four main layers, including data collection, data preprocessing, data processing, and data aggregation and sharing. To overcome the trust issue of crowdsourcing systems, we have proposed a malicious user detection scheme, which identifies anomalies and malicious users. In this algorithm, statistical methods were used to identify outlier data and malicious users. Besides, the user reputation scoring algorithm has been proposed to score and rank unmalicious users. As a result, when rescue teams require additional information, it is possible to communicate with users who have higher scores. Finally, the information aggregation algorithm is responsible for integrating the data collected from unmalicious users. We developed two simulators using MATLAB to evaluate the malicious user detection and user scoring algorithms. The simulation was repeated 200 times, and the average calculation was taken into account. The first simulation has shown that the proposed solution accurately identified outliers and malicious users. It should be noted that we set the threshold for identifying malicious users according to repeated simulation executions. Although this threshold is expected to act satisfactorily in practical real scenarios, it might not be the best threshold value for those scenarios and could be refined. Besides, the second simulation has shown that users who provided information in low-participated areas received higher scores and were more valuable to relief teams, which validates the proposed scoring scheme. The overall aggregated simulation results fulfilled our expectations well. However, one might devise some exceptional cases in which the simulation results, for one instance, become defective.

The proposed crowdsourcing system offers two major advantages compared to similar works. Firstly, it employs a malicious user detection algorithm to ensure the accuracy of gathered data, thus providing reliable results for rescue forces. Besides, this system endeavors to minimize the time required for malicious user detection through the use of statistical methods. Second, unlike related research, this study's user reputation scoring algorithm takes into account not only behavioral parameters but also training and hardware parameters. This incorporation of behavioral, training, and hardware parameters enhances the accuracy of the reputation score calculation. The proposed scoring algorithm encourages users to upload files, which could provide valuable information to the rescue organizations. However, managing a vast number of files in the situation of contributing many people could impose bottlenecks to the server that would require cloud computing. Investigating cloud computing for proposing file management methods is a future research direction. It would encompass image, voice, and video processing techniques for extracting required information from submitted files. Finally, utilizing natural language processing or generative artificial intelligence can be valuable in analyzing and processing the responses to descriptive questions. These technologies can help automate the examination of the answers for extracting required information, which saves time and resources for the relief teams.

Statements and Declarations

Competing interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Ethical Approval

Not applicable

Funding

Not applicable

Authorship Statement

Sajedeh Abbasi: Conceptualization, Software, Original draft preparation

Hamed Vahdat-Nejad: -Supervision- Conceptualization, Methodology- Reviewing and Editing

Hossein Moradi: Supervision- Reviewing and Editing

Computer Code Availability

[HTTPS://GITHUB.COM/SAJEDEH2112/CROWDSOURCING-FOR-FLOOD-DISASTER-EVALUATION](https://github.com/sajedeh2112/crowdsourcing-for-flood-disaster-evaluation)

Open source, 2023, Developed in Python

References

Abbasi, S., Vahdat-Nejad, H., & Hajiabadi, H. (2022). *Trustable Mobile Crowd Sourcing for Acquiring Information from a Flooded Smart Area.* Paper presented at the Smart Cities, Internet of Things and Applications, Iran.

Abrahams, J. (2001). Disaster management in Australia: The national emergency management system. *Emergency Medicine, 13*, 165-173.

Allahbakhsh, M., Ignjatovic, A., Benatallah, B., Bertino, E., & Foo, N. (2012). *Reputation management in crowdsourcing systems.* Paper presented at the International conference on collaborative computing: networking, applications and worksharing, Pittsburgh, PA, USA.

Bahadori, H., Vahdat-Nejad, H., & Moradi, H. (2022). CrowdBIG: crowd-based system for information gathering from the earthquake environment. *Natural Hazards, 114*(3), 3719-3741.

Burkard, S., Fuchs-Kittowski, F., & de Bhroithe, A. O. F. (2017). *Mobile crowd sensing of water level to improve flood forecasting in small drainage areas.* Paper presented at the Environmental Software Systems. Computer Science for Environmental Protection, Croatia.

Caballero-Anthony, M., Cook, A. D., & Chen, C. (2021). Knowledge management and humanitarian organisations in the Asia-Pacific: Practices, challenges, and future pathways. *International Journal of Disaster Risk Reduction, 53*, 102007.

Chen, M., Yang, J., Zhu, X., Wang, X., Liu, M., & Song, J. (2017). Smart home 2.0: Innovative smart home system powered by botanical IoT and emotion detection. *Mobile Networks and Applications, 22*(6), 1159-1169.

Cronstedt, M. (2002). Prevention, preparedness, response, recovery-an outdated concept? *Australian Journal of Emergency Management, The, 17*, 10-13.

Eckhardt, D., Leiras, A., & Thomé, A. M. T. (2022). Using Social Media for Economic Disaster Evaluation: A Systematic Literature Review and Real Case Application. *Natural Hazards Review, 23*, 05021020.

Estellés-Arolas, E., & González-Ladrón-de-Guevara, F. (2012). Towards an integrated crowdsourcing definition. *Journal of Information science, 38*, 189-200.

Fettke, P., & Loos, P. (2003). Classification of reference models: a methodology and its application. *Information systems and e-business management, 1*, 35-53.

Fienen, M. N., & Lowry, C. S. (2012). Social. Water—A crowdsourcing tool for environmental data acquisition. *Computers & Geosciences, 49*, 164-169.

Frigerio, S., Schenato, L., Bossi, G., Mantovani, M., Marcato, G., & Pasuto, A. (2018). Hands-on experience of crowdsourcing for flood risks. An android mobile application tested in Frederikssund, Denmark. *International Journal of Environmental Research and Public Health, 15*, 1926.

Gao, H., Barbier, G., & Goolsby, R. (2011). Harnessing the crowdsourcing power of social media for disaster relief. *IEEE Intelligent Systems, 26*, 10-14.

Golumbic, Y. N., Scroggie, K. R., Kenneally, C. R., Lin, J., Blyth, M. T., Firmer, G., . . . Motion, A. (2023). Meet the Medicines—A Crowdsourced Approach to Collecting and Communicating Information about Essential Medicines Online. *International Journal of Environmental Research and Public Health, 20*, 4242.

Goolsby, R. (2010). Social media as crisis platform: The future of community maps/crisis maps. *ACM Transactions on Intelligent Systems and Technology (TIST), 1*, 1-11.

Hultquist, C., & Cervone, G. (2020). Integration of crowdsourced images, USGS networks, remote sensing, and a model to assess flood depth during Hurricane Florence. *Remote Sensing, 12*, 834.

Kaleem, A., Majeed, A., Khan, T. A., Afzal, H., & Bashir, F. (2015). *Volunteer Reputation evaluation for emergency response operations.* Paper presented at the International Conference on Information and Communication Technologies for Disaster Management, Rennes, France.

Kodheli, O., Lagunas, E., Maturo, N., Sharma, S. K., Shankar, B., Montoya, J. F. M., . . . Kisseleff, S. (2020). Satellite communications in the new space era: A survey and future challenges. *IEEE Communications Surveys & Tutorials, 23*, 70-109.

Kohler, T. (2015). Crowdsourcing-based business models: How to create and capture value. *California management review, 57*, 63-84.

Liu, Y., Piyawongwisal, P., Handa, S., Yu, L., Xu, Y., & Samuel, A. (2011). *Going beyond citizen data collection with mapster: a mobile+ cloud real-time citizen science experiment.* Paper presented at the Seventh International Conference on E-Science Workshops, Sweden.

Ludwig, T., Siebigteroth, T., & Pipek, V. (2014). *Crowdmonitor: Monitoring physical and digital activities of citizens during emergencies.* Paper presented at the International Conference on Social Informatics, Barcelona, Spain.

Ludwig, T., Siebigteroth, T., & Pipek, V. (2015). *Crowdmonitor: Monitoring physical and digital activities of citizens during emergencies.* Paper presented at the SocInfo 2014 International Workshops, Spain.

Mao, K., Capra, L., Harman, M., & Jia, Y. (2017). A survey of the use of crowdsourcing in software engineering. *Journal of Systems and Software, 126*, 57-84.

Naik, N. (2016). *Flooded streets—A crowdsourced sensing system for disaster response: A case study.* Paper presented at the International Symposium on Systems Engineering, UK.

Neumayer, E., Plümper, T., & Barthel, F. (2014). The political economy of natural disaster damage. *Global Environmental Change, 24*, 8-19.

Noorian, Z., & Ulieru, M. (2010). The state of the art in trust and reputation systems: a framework for comparison. *Journal of theoretical and applied electronic commerce research, 5*, 97-117.

Ramesh, M. V., Sudarshan, V., Harilal, G. T., Singh, B., Sudheer, A., & Ekkirala, H. C. (2022). In *Civil Engineering for Disaster Risk Reduction* (the first ed.).

Rossi, C., Stemberger, W., Bielski, C., Zeug, G., Costa, N., Poletto, D., . . . Dominici, F. (2015). *Coupling crowdsourcing, earth observations, and e-gnss in a novel flood emergency service in the cloud.* Paper presented at the International Geoscience and Remote Sensing Symposium Italy.

Sahay, A., Kumar, A. A., Pongpaichet, S., & Jain, R. (2017). *Multimedia rescue systems for floods.* Paper presented at the Proceedings of the 9th International Conference on Management of Digital EcoSystems, Bangkok Thailand.

Schnebele, E., Cervone, G., & Waters, N. (2014). Road assessment after flood events using non-authoritative data. *Natural Hazards and Earth System Sciences, 14*, 1007-1015.

See, L. (2019). A review of citizen science and crowdsourcing in applications of pluvial flooding. *Frontiers in Earth Science, 7*, 44.

Sermet, Y., Villanueva, P., Sit, M. A., & Demir, I. (2020). Crowdsourced approaches for stage measurements at ungauged locations using smartphones. *Hydrological Sciences Journal, 65*, 813-822.

Shi, B., Zhao, J., & Chen, P.-J. (2017). Exploring urban tourism crowding in Shanghai via crowdsourcing geospatial data. *Current Issues in Tourism, 20*, 1186-1209.

Sievers, J. A. (2015). Embracing crowdsourcing: A strategy for state and local governments approaching "whole community" emergency planning. *State and Local Government Review, 47*, 57-67.

Suri, N., Zielinski, Z., Tortonesi, M., Fuchs, C., Pradhan, M., Wrona, K., . . . Pellegrini, V. (2018). *Exploiting smart city IoT for disaster recovery operations.* Paper presented at the 2018 IEEE 4th World Forum on Internet of Things (WF-IoT), Singapore.

Tucker, J. D., Day, S., Tang, W., & Bayus, B. (2019). Crowdsourcing in medical research: concepts and applications. *PeerJ, 7*, e6762.

Vadavalli, A., & Subhashini, R. (2023). A novel truth prediction algorithm for ascertaining the truthfulness of the data and reliability of the users in crowdsourcing application. *Soft Computing, 27*, 1685-1698.

Vahdat-Nejad, H., Asani, E., Mahmoodian, Z., & Mohseni, M. H. (2019). Context-aware computing for mobile crowd sensing: A survey. *Future Generation Computer Systems, 99*, 321-332.

Vahdat-Nejad, H., Bahadori, H., & Abiri, A. (2021). *Information Gathering of Earthquake Disasters by Mobile Crowd Sourcing in Smart Cities.* Paper presented at the 2021 5th International Conference on Internet of Things and Applications (IoT), Isfahan, Iran.

Victorino, J. N. C., Estuar, M. R. J. E., & Lagmay, A. M. F. A. (2016). *Validating the voice of the crowd during disasters.* Paper presented at the International Conference on Social Computing, Behavioral-Cultural Modeling and Prediction and Behavior Representation in Modeling and Simulation, Washington, USA.

Wang, R.-Q., Mao, H., Wang, Y., Rae, C., & Shaw, W. (2018). Hyper-resolution monitoring of urban flooding with social media and crowdsourcing data. *Computers & Geosciences, 111*, 139-147.

Witherow, M. A., Elbakary, M. I., Iftekharuddin, K. M., & Cetin, M. (2017). *Analysis of crowdsourced images for flooding detection.* Paper presented at the European Congress on Computational Methods in Applied Sciences and Engineering, Porto, Portugal.

Wu, G., Zhou, L., Xia, J., Li, L., Bao, X., & Wu, X. (2023). Crowdsourcing Truth Inference Based on Label Confidence Clustering. *ACM Transactions on Knowledge Discovery from Data, 17*, 1-20.

Yang, G., He, S., & Shi, Z. (2016). Leveraging crowdsourcing for efficient malicious users detection in large-scale social networks. *IEEE Internet of Things Journal, 4*, 330-339.

Yu, H., Shen, Z., Miao, C., & An, B. (2012). *Challenges and opportunities for trust management in crowdsourcing.* Paper presented at the Web Intelligence and Intelligent Agent Technology, Macau, China.

Yu, H., Shen, Z., Miao, C., Leung, C., & Niyato, D. (2010). A survey of trust and reputation management systems in wireless communications. *Proceedings of the IEEE, 98*, 1755-1772.

Zhao, S., Pan, G., Zhao, Y., Tao, J., Chen, J., Li, S., & Wu, Z. (2016). Mining user attributes using large-scale app lists of smartphones. *IEEE Systems Journal, 11*, 315-323.

Zuo, Y., Yue, M., Zhang, M., Li, S., Ni, S., & Yuan, X. (2023). OFDM-Based Massive Connectivity for LEO Satellite Internet of Things. *IEEE Transactions on Wireless Communications, 20*, 1-1.