# Disaster Management and Information Transmission Decision-making in Public Safety Systems

Georgios Fragkos*, Eirini Eleni Tsiropoulou*, and Symeon Papavassiliou‖
{gfragkos@unm.edu, eirini@unm.edu, papavass@mail.ntua.gr}
* *Dept. of Electrical and Computer Engineering, The University of New Mexico*, Albuquerque, NM, USA
‖ *School of Electrical and Computer Engineering, National Technical University of Athens*, Athens, Greece

***Abstract*—This paper introduces a Multi-Agency DisAster Management (MADAM) framework for Unmanned Aerial Vehicle (UAV)-assisted public safety systems, based on the principles of game theory and reinforcement learning. Initially, the information quality and criticality (IQC) provided by each agency to an UAV-assisted public safety network is introduced and quantified, and the concept of Value of Information (VoI) that measures each agency's positive contribution to the overall disaster management process is defined. Based on these, a holistic cost function is adopted by each agency, reflecting its relative abstention from the information provisioning process. Each agency aims at minimizing its personal cost function in order to better contribute to the disaster management. This optimization problem is formulated as a non-cooperative game among the agencies and it is proven to be an exact potential game, thus guaranteeing the existence of at least one Pure Nash Equilibrium (PNE). We propose a binary log-linear reinforcement learning algorithm that converges to the optimal PNE. The performance of the proposed approach is evaluated through modeling and simulation under several scenarios, and its superiority compared to other approaches is demonstrated.**

***Index Terms*—Disaster Management, Unmanned Aerial Vehicle, Information Quality and Criticality, Game Theory, Reinforcement Learning.**

## I. Introduction

In public safety events, communication plays an important role throughout the overall disaster management operation. Various agencies are involved in the disaster response operations, e.g., Non-Governmental Organizations, relief and government agencies, reporting their collected information to the Emergency Operation Center (EOC), which is responsible for the disaster response planning. In the traditional Public Safety Networks (PSNs), the agencies' communication with the EOC was enabled by the TETRA or Project 25, which use exclusive bands and specialized hardware, while offer low data rate [1].

Recent additions to the PSN era include broadband Long Term Evolution (LTE) technology [1]. A further addition to improve the communication of the agencies with the EOC is the use of Unmanned Aerial Vehicles (UAVs) acting as relays, overcoming the problem of the damaged communication infrastructure [2]. Extensive research efforts in the relevant recent literature have been devoted to: (i) the optimal resource management in the UAV-assisted PSNs [3], (ii) the optimal UAVs positioning in the disaster area [4], and (iii) the coordination of the LTE undamaged /remaining infrastructure, the UAV-base stations and the WiFi access points [5].

Simply guaranteeing the communication between the disaster relief agencies and the EOC, though critical, is not sufficient to achieve a successful disaster response planning, as the EOC has to cope with incomplete information. Thus, the Information Quality (IQ) is essential for the EOC, since partial information can have catastrophic impact on both the rescue teams and the victims [6]. Various metrics have been proposed in the literature to capture the IQ in PSNs, which can be summarized as follows: (a) product level parameters (e.g., correctness), (b) community level metrics (e.g., usefulness), (c) contextual IQ (e.g., timeliness, amount of information), (d) infrastructure level parameters (e.g., security, response time), and (e) process level metrics (e.g., traceability, interactivity) [7]. The role of IQ in each one of the aforementioned research pillars relevant to the use of UAV in PSNs, i.e., (i)-(iii), has been confronted by several research efforts in an isolated manner though. Thus the unprecedented need arises to develop a holistic approach to enable efficient and effective autonomous multi-agency disaster management, while accounting for the quality of the collected information.

### A. Contributions & Outline

In this paper, our goal is to tackle exactly this problem by introducing a holistic distributed framework supporting the multi-agencies' self-optimization in terms of reporting their information to a UAV that hovers over the disaster-struck area, and particularly enables the communication of the various agencies with the EOC due to the damaged ground infrastructure. The main technical contributions of this research work are summarized as follows.

1. A multi-agency disaster management system is considered, which is supported by an UAV enabling the communication between the agencies and the EOC by acting as a relay. The novel concept of Information Quality and Criticality (IQC) is introduced to quantify the importance level of the agencies' transmitted information. Combining the IQC with the amount of the provided information by each agency, we introduce the concept of Value of Infor-

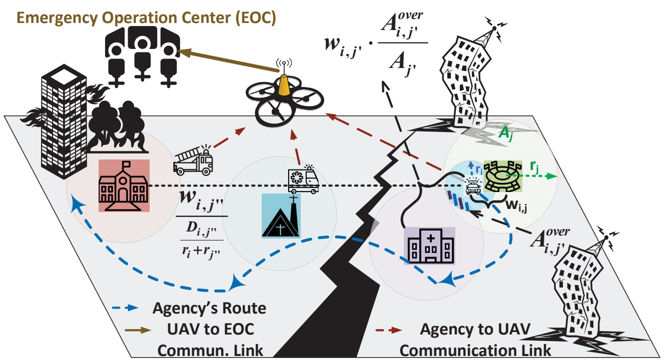


Fig. 1: Multi-agency Disaster Management Topology

mation (VoI), which captures each agency's contribution to the disaster management operation (Section II).
2. Each agency is characterized by a cost function that captures its relative abstention from the information provisioning process to the EOC compared to the total information contribution by all the agencies based on the IQC and VoI. In this setting, each agency aims at minimizing its experienced cost in order to efficiently contribute to the disaster management process (Section III.A). The minimization problem of each agency's cost function is formulated as a non-cooperative game, while we prove that it is an exact potential game, and thus there exists at least one Pure Nash Equilibrium (PNE) (Section III.B).
3. To determine the PNE, the agency's cost minimization problem is transformed to the corresponding maximization problem of the formulated game's potential function. This problem is solved by appropriately adapting the binary log-linear reinforcement learning algorithm, referred to as B-logit [8]. The B-logit algorithm determines among all PNEs, the one that maximizes the introduced potential function. A detailed complexity analysis of the overall Multi-Agency DisAster Management (MADAM) framework is provided that demonstrates its feasibility and applicability in real-life scenarios (Section IV).
4. Detailed numerical results, obtained via modeling and simulation, show that the proposed framework concludes to a promising solution for realizing an effective multi-agency disaster management process, which efficiently scales in large-scale use case scenarios while significantly outperforms other alternatives (Section V). Finally, Section VI concludes the paper.

## II. System Model

A heterogeneous wireless PSN is considered consisting of $|N|$ different agency mobile infrastructure objects, e.g., police cars, fire trucks, ambulances. We consider $|N| \geq 3$ for realistic applications, while $N$ denotes the agencies' set. A set of Points of Interest (POIs) $J = \{1, ..., j, ..., |J|\}$, e.g., schools, religion institutions, stadiums, exists in the disaster area, where the humans tend to gather. A UAV hovers above the disaster area with coordinates $(x_{UAV}, y_{UAV}, z_{UAV})$ and covers a disaster area of radius $R$ (e.g., $R = 1800m$). The UAV acts as a relay to forward the agencies' information to the EOC, which is out of the agencies' communication range while the ground communication infrastructure is assumed to be damaged. The POIs are static, having $(x_j, y_j)$ coordinates on the disaster's area plane. The agencies are moving to collect information from the POIs and help the victims, thus, their respective coordinates on the disaster's plane are time varying. The overall topology is presented in Fig. 1.

### A. Communication Model

The agencies send their data to the UAV, which acts as a relay transmitting the overall collected data to the EOC. The access network (agencies to UAV), is examined in this paper assuming Non-Orthogonal Multiple Access (NOMA), while the backbone network (UAV to EOC), is considered to have sufficient bandwidth in the single link communication to transmit the overall data. The channel gain between the agency $i$ and the UAV is $G_i = \frac{k}{d_i^2}$, where $k$ expresses the shadow effect via a lognormal random variable, and $d_i$ is the distance between the agency $i$ and the UAV. Noting that the power control of the agencies' transmissions is not the key objective of this paper, we consider a simplistic, but realistic, transmission power model, as $P_i = \frac{d_i}{R} P_i^{Max}$, where the agency's transmission power is relative to its distance from the receiver and $P_i^{Max} = 1W$ is the agency's maximum transmission power. Due to the NOMA's Successive Interference Cancellation technique [9], the interference is $I = \sum_{i' \geq i+1}^{|N|} G_{i'} \cdot P_{i'} + I_o$, where $I_o$ is the thermal noise captured as Additive White Gaussian Noise. Based on Shannon's formula, the agency's data rate is $R_i^{(t)} = B \cdot log_2(1 + \frac{P_i \cdot G_i}{I})$, where $B$ is the system's bandwidth. At each time slot $t$, each agency transmits to the UAV a ratio of information $\mathrm{a}_i^{(t)} \in \mathbb{A}_i^{(t)} = \{\mathrm{a}_i^{min^{(t)}}, \ldots, \mathrm{a}_i^{max^{(t)}}\}, |\mathbb{A}_i^{(t)}| \in \mathbb{N}$ ($|\mathbb{A}_i^{(t)}|$: cardinality of $\mathbb{A}_i^{(t)}$) from its total amount of information $I_i^{Max(t)}$. The latter stems from the argument that the transmission of whole $I_i^{Max(t)}$ is neither always useful for the EOC, nor efficient due to bandwidth limitations. The amount of agency's $i$ transmitted information at time slot $t$ is given as follows.

$$f_i^{(t)}(\mathrm{a}_i^{(t)}) = (\mathrm{a}_i^{(t)} \cdot I_i^{Max(t)}) \cdot R_i^{(t)} \tag{1}$$

### B. Information Quality & Criticality

Let us denote by $IQC_{i,j}^{(t)}$ the Information Quality and Criticality of each agency $i$ with respect to each PoI $j$. This metric reflects the information's importance, relevance, as well as the thematic, temporal and positional accuracy of the reported information, where all those different angles are comprehensively captured by a weighting factor $w_{i,j}$, $w_{i,j} \in (0, 1]$. Each PoI and each agency have an information gathering area of radius $r_j$ and $r_i$, $r_j > r_i$, respectively. Thus, if the agencies are within the information gathering area of the PoI, then they may report information with relatively high $IQC_{i,j}^{(t)}$, taken into account that the information quality is improved if agencies collect information closer from the field of a given PoI where an incident occurred. Therefore, the $IQC_{i,j}^{(t)}$ of each agency $i$ with respect to each PoI $j$ is defined as follows:

$$IQC_{i,j}^{(t)} = \begin{cases} w_{i,j}, & r_j \geq D_{i,j} + r_i \\ w_{i,j} \cdot \frac{\mathrm{A}_{i,j}^{over}}{\mathrm{A}_j}, & r_j < D_{i,j} + r_i \; and \; r_i + r_j > D_{i,j} \\ \frac{w_{i,j}}{\frac{D_{i,j}}{r_i + r_j}}, & r_j + r_i \leq D_{i,j} \end{cases} \tag{2}$$

where $D_{i,j}$ is the Euclidean distance between the centers of the agency's $i$ and the PoI's $j$ information gathering areas, $A_j = \pi \cdot r_j^2$ is the information gathering area characterizing the PoI $j$ and $A_{i,j}^{over} = r_i^2 \cdot cos^{-1}(\frac{D_{i,j}^2 + r_i^2 - r_j^2}{2 \cdot D_{i,j} \cdot r_i} + r_j^2 \cdot cos^{-1}(\frac{D_{i,j}^2 + r_j^2 - r_i^2}{2 \cdot D_{i,j} \cdot r_j}) - \frac{1}{2}\sqrt{(-D_{i,j} + r_j + r_i) \cdot (D_{i,j} + r_i - r_j) \cdot (D_{i,j} - r_i + r_j)} \cdot \sqrt{(D_{i,j} + r_i + r_j)}$ is the overlapping area between the two information gathering areas, i.e., $A_i, A_j$ (see Fig. 1). The average IQC of each agency, reflecting the importance of the agency's $i$ reported information regarding all the PoIs, is obtained as: $IQC_i^{(t)} = \frac{\sum_{j=1}^{|J|} IQC_{i,j}^{(t)}}{|J|}$. Combining the relative amount of information $f_i^{(t)}(\mathrm{a}_i^{(t)})$ that each agency reports and the $IQC_i^{(t)}$, we define the value of information $VoI_i^{(t)}$ of agency $i$ per time slot $t$ as follows.

$$VoI_i^{(t)}(\mathrm{a}_i^{(t)}, \mathbf{a}_{-\mathbf{i}}^{(\mathbf{t})}) = \frac{f_i^{(t)}(\mathrm{a}_i^{(t)})}{\sum_{\mathrm{k} \neq i, \mathrm{k} \in N} f_{\mathrm{k}}^{(t)}(\mathrm{a}_{\mathrm{k}}^{(t)})} \cdot IQC_i^{(t)} \tag{3}$$

Also, the Value of Information that each agency provides to the EOC should be evaluated over the past time. Thus, we define the relative value of information of each agency with respect to the others over the past time as follows.

$$\hat{VoI}_i^{(t)}(\mathrm{a}_i^{(t)}, \mathbf{a}_{-\mathbf{i}}^{(\mathbf{t})}) = \frac{\sum_{t' \in [1, t-1]} VoI_i^{(t')}(\mathrm{a}_i^{(t')}, \mathbf{a}_{-\mathbf{i}}^{(\mathbf{t}')})}{\sum_{t' \in [1, t-1]} \sum_{\mathrm{k} \neq i, \mathrm{k} \in N} VoI_{\mathrm{k}}^{(t')}(\mathrm{a}_{\mathrm{k}}^{(t')}, \mathbf{a}_{-\mathbf{k}}^{(\mathbf{t}')})} \tag{4}$$

where $t$ denotes the current time slot. For the special case of $t = 1$, we assume that $\hat{VoI}_i^{(t)}$ is calculated based on previous experience of the agencies from past events. In the following, for notational convenience we have $f_i^{(t)} = f_i^{(t)}(\mathrm{a}_i^{(t)})$ and $\hat{VoI}_i^{(t)} = \hat{VoI}_i^{(t)}(\mathrm{a}_i^{(t)}, \mathbf{a}_{-\mathbf{i}}^{(\mathbf{t})})$.

## III. Multi-agency Disaster Management

### A. Agency's Cost Function

Each agency $i$ is associated with a cost function (properly formulated in Eq.5) representing its relative abstention from contributing information to the EOC compared to the other agencies (numerator of Eq.5) and compared to the provided information by all the agencies in the previous time slot (denominator of Eq.5). Each agency aims at minimizing its cost function (Eq. 5) towards significantly contributing to the disaster management.

$$C_i^{(t)}(\mathrm{a}_i^{(t)}, \mathbf{a}_{-\mathbf{i}}^{(\mathbf{t})}) = \frac{\sum_{\mathrm{k} \in N, \mathrm{k} \neq i} (f_{\mathrm{k}}^{(t)} V\hat{o}I_{\mathrm{k}}^{(t)}) - f_i^{(t)} V\hat{o}I_i^{(t)}}{\sum_{\mathrm{k} \in N} f_{\mathrm{k}}^{(t-1)} V\hat{o}I_{\mathrm{k}}^{(t-1)}} \tag{5}$$

Given the the partial information availability among the agencies, a distributed minimization problem of each agency's cost function is formulated as follows.

$$\min_{\mathrm{a}_i^{(t)} \in \mathbb{A}_i^{(t)}} C_i^{(t)}(\mathrm{a}_i^{(t)}, \mathbf{a}_{-\mathbf{i}}^{(\mathbf{t})}) \tag{6}$$

### B. Exact Potential Game

Based on the distributed nature of the minimization problem (6), it is confronted as a non-cooperative game $G = [N, \{\mathbb{A}_i^{(t)}\}_{i \in N}, \{C_i^{(t)}\}_{i \in N}]$, where $N$ is the set of agencies, $\mathbb{A}_i^{(t)}$ is the set of agency's $i$ strategies at time slot $t$, and $C_i^{(t)}$ is the agency's $i$ cost function. The solution of the game $G$ is a Pure Nash Equilibrium (PNE) regarding all the agencies' information transmission ratios.

*Definition 1:* An information ratios vector $\mathbf{a}^{*(\mathbf{t})}$ in the strategy set $\mathbb{A}^{(t)} = \times_{\forall i \in N} \mathbb{A}_i^{(t)}$ is a PNE of the game $G$ if $\forall \mathrm{a}_i^{(t)}, \forall i \in N$, it holds true that $C_i^{(t)}(\mathrm{a}_i^{*(t)}, \mathbf{a}_{-\mathbf{i}}^{*(\mathbf{t})}) \leq C_i^{(t)}(\mathrm{a}_i^{(t)}, \mathbf{a}_{-\mathbf{i}}^{*(\mathbf{t})})$.

Towards ensuring the existence of a PNE for the game $G$, we prove that it is an exact potential game. An exact potential game is characterized by its exact potential function $\Phi(\mathrm{a}_i^{(t)}, \mathbf{a}_{-\mathbf{i}}^{(\mathbf{t})})$) that exactly reflects any unilateral change in each agency's cost function, as follows.

$$C_i^{(t)}(\mathrm{a}_i^{(t)}, \mathbf{a}_{-\mathbf{i}}^{(\mathbf{t})}) - C_i^{(t)}(\mathrm{a}_i^{'(t)}, \mathbf{a}_{-\mathbf{i}}^{(\mathbf{t})}) = \Phi(\mathrm{a}_i^{(t)}, \mathbf{a}_{-\mathbf{i}}^{(\mathbf{t})}) - \Phi(\mathrm{a}_i^{'(t)}, \mathbf{a}_{-\mathbf{i}}^{(\mathbf{t})}) \tag{7}$$

where $\mathrm{a}_i^{(t)}, \mathrm{a}_i^{'(t)} \in \mathbb{A}_i^{(t)}, \mathrm{a}_i^{(t)} \neq \mathrm{a}_i^{'(t)}$.

*Theorem 1:* The non-cooperative game $G$ is an exact potential game with an exact potential function given by Eq.8 and has at least one PNE $\mathbf{a}^{*(\mathbf{t})} \in \mathbb{A}^{(t)}$.

$$\Phi(\mathrm{a}_i^{(t)}, \mathbf{a}_{-\mathbf{i}}^{(\mathbf{t})}) = \sum_{i \in N} \frac{f_i^{(t)} V\hat{o}I_i^{(t)} - \sum_{\mathrm{k} \in N, \mathrm{k} \neq i} f_{\mathrm{k}}^{(t)} V\hat{o}I_{\mathrm{k}}^{(t)}}{(|N| - 2) \sum_{\mathrm{k} \in N} f_{\mathrm{k}}^{t-1} V\hat{o}I_{\mathrm{k}}^{(t-1)}} \tag{8}$$

$$\begin{aligned} \Phi(\mathrm{a}_i^{(t)}, \mathbf{a}_{-\mathbf{i}}^{(\mathbf{t})}) &= \frac{\sum_{i \in N} [f_i^{(t)} V\hat{o}I_i^{(t)} - (|N| - 1) f_i^{(t)} V\hat{o}I_i^{(t)}]}{(|N| - 2) \sum_{\mathrm{k} \in N} f_{\mathrm{k}}^{(t-1)} V\hat{o}I_{\mathrm{k}}^{(t-1)}} \\ &= -\frac{\sum_{i \in N} f_i^{(t)} V\hat{o}I_i^{(t)}}{\sum_{\mathrm{k} \in N} f_{\mathrm{k}}^{(t-1)} V\hat{o}I_{\mathrm{k}}^{(t-1)}} \end{aligned} \tag{9}$$

*Proof:* Towards proving Theorem 1, the condition of Eq.7 should hold true $\forall i \in N$. The exact potential function (Eq.8) can be written as presented in Eq. 9. Let $\mathrm{a}_i^{'} \in \mathbb{A}_i, \mathrm{a}_i^{'} \neq \mathrm{a}_i$, then, the outcome of Eq. 10 is derived, given that $V\hat{o}I_i^{(t)}(\mathrm{a}_i, \mathbf{a}_{-\mathbf{i}})$ depends only on agency's $i$ strategy $\mathrm{a}_i$ in the previous time slots $[1, t-1]$. Thus, we conclude that Eq.7 holds true. Therefore, the game $G$ is an exact potential game and at least one PNE exists [10].

$$
\begin{aligned}
&\Phi(\mathrm{a}_i^{(t)}, \mathbf{a}_{-\mathbf{i}}^{(\mathbf{t})}) - \Phi(\mathrm{a}_i^{'(t)}, \mathbf{a}_{-\mathbf{i}}^{(\mathbf{t})}) \\
&= -\frac{\sum_{i \in N} f_i^{(t)}(\mathrm{a}_i^{(t)}) \cdot \hat{VoI_i}^{(t)}(\mathrm{a}_i^{(t)}, \mathbf{a}_{-\mathbf{i}}^{(\mathbf{t})})}{\sum_{\mathrm{k} \in N} f_{\mathrm{k}}^{(t-1)}(\mathrm{a}_{\mathrm{k}}^{(t-1)}) \cdot \hat{VoI}_{\mathrm{k}}^{(t-1)}(\mathrm{a}_{\mathrm{k}}^{(t-1)}, \mathbf{a}_{-\mathbf{k}}^{(\mathbf{t-1})})} \\
&+ \frac{\sum_{i \in N} f_i^{(t)}(\mathrm{a}_i^{'(t)}) \cdot \hat{VoI_i}^{(t)}(\mathrm{a}_i^{'(t)}, \mathbf{a}_{-\mathbf{i}}^{(\mathbf{t})})}{\sum_{\mathrm{k} \in N} f_{\mathrm{k}}^{(t-1)}(\mathrm{a}_{\mathrm{k}}^{(t-1)}) \cdot \hat{VoI}_{\mathrm{k}}^{(t-1)}(\mathrm{a}_{\mathrm{k}}^{(t-1)}, \mathbf{a}_{-\mathbf{k}}^{(\mathbf{t-1})})} \qquad (10) \\
&= \frac{f_i^{(t)}(\mathrm{a}_i^{'(t)}) \cdot \hat{VoI_i}^{(t)}(\mathrm{a}_i^{'(t)}, \mathbf{a}_{-\mathbf{i}}^{(\mathbf{t})}) - f_i^{(t)}(\mathrm{a}_i^{(t)}) \cdot \hat{VoI_i}^{(t)}(\mathrm{a}_i^{(t)}, \mathbf{a}_{-\mathbf{i}}^{(\mathbf{t})})}{\sum_{\mathrm{k} \in N} f_{\mathrm{k}}^{(t-1)}(\mathrm{a}_{\mathrm{k}}^{(t-1)}) \cdot \hat{VoI}_{\mathrm{k}}^{(t-1)}(\mathrm{a}_{k}^{(t-1)}, \mathbf{a}_{-\mathbf{k}}^{(\mathbf{t-1})})} \\
&= C_i(\mathrm{a}_i^{(t)}, \mathbf{a}_{-\mathbf{i}}^{(\mathbf{t})}) - C_i(\mathrm{a}_i^{'(t)}, \mathbf{a}_{-\mathbf{i}}^{(\mathbf{t})})
\end{aligned}
$$

## IV. Disaster Management Planning based on Reinforcement Learning

In this section, we propose the Multi-Agency DisAster Management (MADAM) algorithm based on the principles of the binary log-linear reinforcement learning algorithm, called B-logit. The MADAM algorithm requires no information to be exchanged among the agencies and converges to the optimal PNE of the game $G$, in the sense that maximizes the potential function (Eq.8) and equivalently minimizes each agency's cost function (Eq.5) [8]. Initially, each agency selects an information transmission ratio $\mathrm{a}_i^{(ite=0)^{(t)}}$ with equal probability $Pr(\mathrm{a}_i^{(ite=0)^{(t)}}) = 1/|\mathbb{A}_i^{(t)}|$. At each iteration, one agency, based on its priority ID that is assigned offline by the EOC, performs exploration and learning. Thus, at the $ite$ iteration, the agency $i$ randomly tries an alternative strategy $\mathrm{a}_i^{'(ite)^{(t)}}$ with equal probability $1/|\mathbb{A}_i^{(t)}|$, concluding to a cost $C_i^{'(ite)^{(t)}}$ (exploration phase). Let $\mathrm{a}_i^{(ite-1)^{(t)}}$ and $C_i^{(ite-1)^{(t)}}$ denote the selected strategy of the agency $i$ and the cost function at $(ite-1)$ iteration, respectively. At the $ite$ iteration, the agency $i$ updates its strategy according to the following rule, while the rest agencies keep their previous strategies (learning phase).

$$
Pr(\mathrm{a}_i^{(ite)^{(t)}} = \mathrm{a}_i^{'(ite)^{(t)}}) = \frac{e^{-C_i^{'(ite)^{(t)}} \cdot \beta}}{e^{-C_i^{(ite-1)^{(t)}} \cdot \beta} + e^{-C_i^{'(ite)^{(t)}} \cdot \beta}} \qquad (11a)
$$

$$
Pr(\mathrm{a}_i^{(ite)^{(t)}} = \mathrm{a}_i^{(ite-1)^{(t)}}) = \frac{e^{-C_i^{(ite-1)^{(t)}} \cdot \beta}}{e^{-C_i^{(ite-1)^{(t)}} \cdot \beta} + e^{-C_i^{'(ite)^{(t)}} \cdot \beta}} \qquad (11b)
$$

where $\beta$ is the learning parameter. The detailed steps of MADAM algorithm are summarized in Algorithm 1.

Regarding the MADAM algorithm's complexity, for a specific time-slot $t$ we initially determine the agencies' coordinates $(x_i, y_i)$ and the $IQC_i^{(t)}, \forall i \in N$, which concludes to a complexity of $\mathcal{O}(|N|)$. Since at each $ite^{(t)}$ only one agency performs the exploration, and since all the actions $\mathrm{a}_i^{(ite)^{(t)}}, \forall i \in N$, involve only algebraic calculations, i.e., $\mathcal{O}(1)$, the complexity of the MADAM algorithm for a time-slot $t$ is $\mathcal{O}(|N| + Ite^{(t)})$, where $Ite^{(t)}$ is the total number of iterations for convergence in time-slot t. Since $Ite^{(t)} >> |N|$ (see Section V), the complexity in time-slot $t$ is $\mathcal{O}(Ite^{(t)})$. The overall complexity of the MADAM Algorithm for all the time-slots $T$ is $\mathcal{O}(T \cdot \max_{t \in [1,T]} \{Ite^{(t)}\})$.

## V. Numerical Results

A detailed numerical evaluation is presented in terms of the overall framework's operation efficiency (Section V-A), its scalability and complexity (Section V-B), and superiority compared to other alternatives (Section V-C). We consider $|N| = 30$, $|J| = 4$, $B = 5\,MHz$, $I_o = 10^{-13}$, $R = 1800m$, $I_i^{Max} \in [150, 250]MB$, $\mathrm{a}_i^{min^{(t)}} \in [0.1, 0.3]$, $\mathrm{a}_i^{max^{(t)}} \in [0.8, 1.0]$ with an intermediate step of 0.1. The agencies follow a random route with velocity $|v_i| \in [6, 9]\frac{m}{sec}$. A detailed Monte Carlo analysis has been executed considering averages over $10,000$ executions.

### *A. Pure Framework Operation Evaluation*

Fig. 2a presents the average cost $C_i$ and the average potential function $\Phi(\mathrm{a}_i, \mathbf{a}_{-\mathbf{i}})$ as a function of MADAM algorithm's iterations till convergence to the PNE (the lower horizontal axis reflects iterations while the upper horizontal axis refers to the actual execution time). We observe that for practical purposes less than 800 iterations, i.e., 0.17 sec, are required to reach the PNE, where the agencies' average abstention from providing information to the EOC is minimized (i.e., $C_i$). Accordingly, as expected from Eq.8, it is also observed that $\Phi$ is maximized.

In Figures 2b and 2c, we examine the detailed behavior of one randomly selected agency throughout a time-slot. Specifically, in Fig. 2b we present the $f_k \cdot \hat{VoI_k}$ and the $\sum_{j \neq k} f_j \cdot \hat{VoI_j}$ (logarithmic scale), as well as the agency's $k$ cost function $C_k$, as a function of the B-logit iterations. As time evolves the specific agency decreases slightly the

**Algorithm 1** MADAM: Multi-Agency DisAster Management Algorithm

- **Input:** $N$, $J$, $w_{i,j}$, $(x_j, y_j)$, *initial* $(x_i, y_i)$, $\forall i \in N$
- **Output:** $\mathbf{a}^{*(\mathbf{t})}$
- **Initialization:** $f_i^{(0)}, \hat{VoI}_i^{(0)}$, $\forall i \in N$, $T$, $d_t = 4sec$, $|v_i| \in [6, 9]\frac{m}{sec}$, $\epsilon = 0.05$
- **for** every time-slot $t \in T$ **do**
  - $ite^{(t)} = 0, Convergence = 0$, *Arbitrary Action Profile* $\mathrm{a}^{(t)}, T'$, *find* $(x_i, y_i)$, $IQC_i^{(t)}, \forall i \in N$
  - **while** $Convergence == 0$ **do**
    - $ite^{(t)} = ite^{(t)} + 1$;
    - *Agency i (identified by priority ID) selects* $\mathrm{a}_i^{'(ite)^{(t)}}$ *with equal probability* $\frac{1}{|\mathbb{A}_i^{(t)}|}$ , *computes* $C_i^{'(ite)^{(t)}}$ *and updates* $\mathrm{a}_i^{(ite)^{(t)}}$ (Eq.11a, 11b)
    - *The other agencies repeat their previous actions, i.e.,* $\mathbf{a}_{-\mathbf{i}}^{(\mathbf{ite})^{(\mathbf{t})}} = \mathbf{a}_{-\mathbf{i}}^{(\mathbf{ite-1})^{(\mathbf{t})}}$
    - **if** $|(\frac{\sum_{ite^{(t)}=0}^{T'} (\Phi^{(ite)^{(t)}})}{T'} - \Phi^{(ite)^{(t)}})| \leq \epsilon$ **then**
      - $Convergence = 1$
    - **end if**
  - **end while**
- **end for**

quantity of the information it aims to send throughout the time-slot ($\hat{VoI}_k$ remains the same during the time-slot), in order to avoid acting in a myopic way. Similarly, all the other agencies reduce their overall $f_j \cdot \hat{VoI}_j$ and eventually agency $k$ experiences a lower cost function $C_k$. In Fig. 2c, we notice that the $k^{th}$ agency's $VoI_k$ value increases over time/iterations, which confirms that the agency learns to send the appropriate amount of information, thus achieving a satisfying combination of quality and quantity of transmitted information.

Fig. 2d illustrates the behavior of the potential function value as a function of the B-logit iterations, for different values of the learning parameter, i.e., $\beta = 100, 500$ and 1000. The results clearly demonstrate a tradeoff between convergence time and optimality. Higher values of $\beta$ conclude to higher values of the potential function $\Phi$ while also presenting increased convergence time. This is due to the fact that for large values of $\beta$ the MADAM algorithm spends more time to optimal states to better explore them.

### B. Scalability and Complexity Evaluation

In this section we adduce an extensive scalability and complexity analysis of the framework performance in terms of increasing number of agencies and granularity of the available number of ratios of information. Fig. 3a presents the following metrics after the MADAM algorithm's convergence (four different curves): (a) overall amount of information transmitted from all the agencies, i.e., $\sum_{i \in N} f_i$, (b) the respective average $f_i$ per agency, (c) the total ratios of information, i.e., $\sum_{i \in N} \mathrm{a}_i$, and (d) the corresponding average $\mathrm{a}_i$ per agency, as the number of the agencies in the coverage area increases.

We observe that even though the summation of all agencies' $\mathrm{a}_i$ increases as $|N|$ increases, the average portion of information per agency (i.e., $\mathrm{a}_i$), has a decreasing trend due to the potential access channel congestion. Moreover, we observe that the total amount of sent information $\sum_{i \in N} f_i$, decreases with respect to the number of the agencies, because each agency's achievable data rate decreases dramatically due to the increased interference. The same holds true for the average $f_i$. In Fig. 3b we present the B-logit execution time required for convergence as a function of the increasing number of agencies. It is noted that, since B-logit is an asynchronous distributed learning algorithm in the sense that in every iteration at a specific time-slot only one agency can perform exploration and update its action, when the number of agencies increases, the corresponding convergence time increases as well.

Subsequently, we examine the behavior of the proposed framework when the agencies are equipped with a potential set of actions of higher granularity, where a larger number of potential alternative actions are available to them. In particular, in Fig. 3c we observe that the average cost function $C_i$ per agency presents a decreasing trend as the number of actions increases, due to the fact that the wider range of available actions allows a more detailed exploration, thus leading to the choice of more efficient strategies. Regarding the execution time, we observe that until a certain number of actions it has a decreasing trend, because the agencies have more alternatives to choose from and the system proceeds faster to convergence. However, after a certain number of additional actions, e.g., 50 in our scenario, the time starts to increase instead, because the agencies spend more time to explore the available strategy space in order to obtain the desired PNE (this is well aligned with the increasing B-logit execution time). Furthermore, in Fig. 3d we present the average $f_i \cdot VoI_i$ per agency as a function of the number of the additional actions to the agencies' strategy space. We observe that due to the agencies' larger strategy space (i.e., number of actions), they can choose more wisely their actions, thus, the combination of the average quantity and of the average quality of information (i.e., $f_i \cdot VoI_i$) that they try to send increases throughout the B-logit iterations .

### C. Comparative Results

In this section, initially, we compare the efficiency of our proposed framework with the following six approaches concerning the agencies' selection of their action. (1) Each agency sends the maximum information. (2) Each agency sends the minimum information. (3) Each agency sends a random portion of its $I_i^{Max(t)}$.(4) Each agency determines its action based only on physical aspects. To realize this the disaster area is divided into different "information zones" around the UAV. The further the agency is from the UAV, the less information gradually it transmits. At the most distant information zone from the UAV, the agency will transmit information $\mathrm{a}_i^{min^{(t)}} \cdot I_i^{Max(t)}$. (5) Each agency determines its action based only on the social aspects $\mathrm{a}_i^{(t)} = \dfrac{\sum_{j \in J} w_{i,j}}{|J|}, \forall i \in N$. (6) Each agency determines its action based on the socio-physical criteria $\mathrm{a}_i^{(t)}{=}IQC_i^{(t)}$. Our examined scenario ran for $200,000$ time-slots.

Fig. 4a depicts the average $VoI_i^{(t)}$ per agency per time-slot, including all the aforementioned alternative strategies. As we can clearly notice, our approach outperforms all the other approaches confirming that efficiently and in a distributed manner orchestrates the agencies to send the appropriate amount of information with the corresponding quality of information. This is followed by the socio-physical approach, being not only aware of the social aspect, but also more adaptive and dynamic (physical aspect), due to the fact that the agency's position changes over time. The approaches that exhibit the worst performance are the ones where the agencies send their maximum ($\mathrm{a}_i^{max^{(t)}}$) or minimum ($\mathrm{a}_i^{min^{(t)}}$) amount of information. This is observed, because in those cases the agencies act in a myopic and static way that leads to a very low average $VoI_i, i \in N$ per agency per time-slot.

Subsequently, we compare our approach, where instead of the B-logit algorithm two alternative algorithms are applied, namely the Max log linear learning algorithm

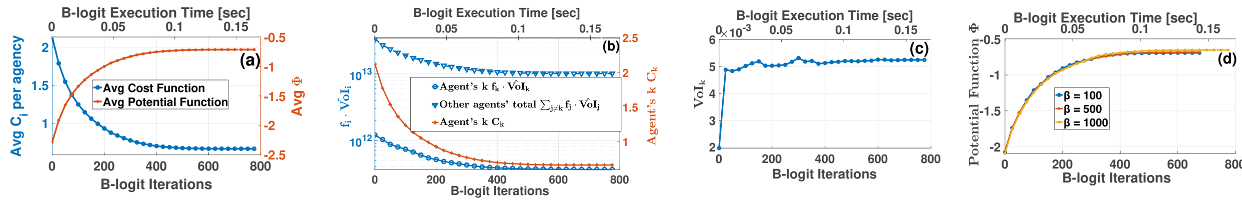


Fig. 2: Pure operation of the proposed framework

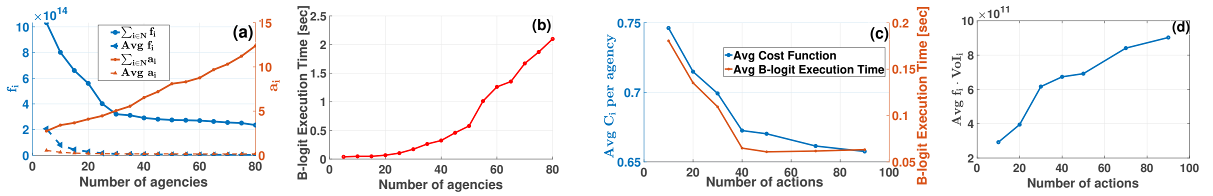


Fig. 3: Scalability and Complexity

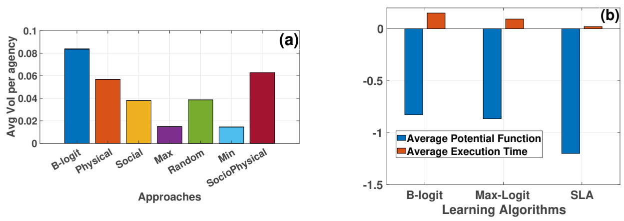


Fig. 4: Comparative Evaluation

(denoted as Max-logit) and the Stochastic Learning Automata (denoted as SLA) [8]. B-logit outperforms both of the other approaches in terms of achieving a better maximization value of the potential function. That happens because B-logit (and Max-logit) converges to the best Nash equilibrium with a high probability, which maximizes the potential function, while the SLA just converges to an arbitrary PNE. For the same reasoning, Max-logit presents performance very close to the B-logit. However, the SLA achieves lower actual execution time, clearly presenting a tradeoff between the efficiency in obtaining the best Nash equilibrium and the algorithm's real execution time. SLA converges faster to an arbitrary PNE due to the fact that it is a synchronous algorithm where in every iteration of a specific time-slot all the agencies can simultaneously perform exploration and change their actions, in contrast to B-logit and Max-logit which are asynchornous in nature.

## VI. Conclusions

In this paper, we propose a multi-agency disaster management framework, where the agencies provide their collected information to the EOC through the UAV, which acts as a relay given that the ground infrastructure is damaged. The concept of Information Quality and Criticality (IQC) is introduced to quantify the importance level of the agency's provided information. Each agency is associated with a holistic cost function, which represents its relative abstention in information provisioning compared to the rest of the agencies. A non-cooperative game is formulated among the agencies and we prove that it is an exact potential game, thus, the existence of at least one PNE is shown. The optimal PNE is determined by the proposed binary log-linear reinforcement learning algorithm. Finally, detailed numerical and comparative results are presented. Part of our current and future work focuses on studying the behavioral characteristics of the agencies towards sharing information with the EOC. Towards this direction we will adopt concepts from the Prospect Theory and the theory of the Tragedy of the Commons.

## Acknowledgment

The research of Dr. Eirini Eleni Tsiropoulou was conducted as part of the NSF CRII-1849739.

## References

[1] A. Prasad, A. Maeder, K. Samdanis, A. Kunz, and G. Velev, "Enabling group communication for public safety in lte-advanced networks," *Journal of Network and Computer Applications*, vol. 62, pp. 41–52, 2016.

[2] A. Merwaday, A. Tuncer, A. Kumbhar, and I. Guvenc, "Improved throughput coverage in natural disasters: Unmanned aerial base stations for public-safety communications," *IEEE Vehicular Technology Magazine*, vol. 11, no. 4, pp. 53–60, 2016.

[3] M. Mozaffari, W. Saad, M. Bennis, and M. Debbah, "Mobile unmanned aerial vehicles (uavs) for energy-efficient internet of things communications," *IEEE Transactions on Wireless Communications*, vol. 16, no. 11, pp. 7574–7589, 2017.

[4] D. Sikeridis, E. E. Tsiropoulou, M. Devetsikiotis, and S. Papavassiliou, "Socio-spatial resource management in wireless powered public safety networks," in *IEEE MILCOM*, 2018, pp. 810–815.

[5] D. Athukoralage, I. Guvenc, W. Saad, and M. Bennis, "Regret based learning for uav assisted lte-u/wifi public safety networks," in *IEEE GLOBECOM*, 2016, pp. 1–7.

[6] J. Lee, N. Bharosa, J. Yang, M. Janssen, and H. R. Rao, "Group value and intention to use - a study of multi-agency disaster management information systems for public safety," *Decision Support Systems*, vol. 50, no. 2, pp. 404–414, 2011.

[7] H. Seppänen and K. Virrantaus, "Shared situational awareness and information quality in disaster management," *Safety Science*, vol. 77, pp. 112–122, 2015.

[8] Y. Xu, J. Wang, and Q. Wu, "Distributed learning of equilibria with incomplete, dynamic, and uncertain information in wireless communication networks," in *Game Theory Framework Applied to Wireless Comm. Net.* IGI Global, 2016, pp. 63–86.

[9] P. Vamvakas, E. E. Tsiropoulou, and S. Papavassiliou, "Dynamic provider selection & power resource management in competitive wireless communication markets," *Mobile Networks and Applications*, vol. 23, no. 1, pp. 86–99, 2018.

[10] E. E. Tsiropoulou, G. K. Katsinis, A. Filios, and S. Papavassiliou, "On the problem of optimal cell selection and uplink power control in open access multi-service two-tier femtocell networks," in *Int. Conf. on Ad-Hoc Networks and Wireless.* Springer, 2014, pp. 114–127.