# Contract-Theoretic Resource Control in Wireless Powered Communication Public Safety Systems

Nathan Patrizi, Georgios Fragkos, Eirini Eleni Tsiropoulou
*Dept. of Electrical and Computer Engineering*
*University of New Mexico*
Albuquerque, NM, USA
npatrizi@unm.edu, gfragkos@unm.edu, eirini@unm.edu

Symeon Papavassiliou
*School of Electrical and Computer Engineering*
*National Technical University of Athens*
Athens, Greece
papavass@mail.ntua.gr

GLOBECOM 2020 - 2020 IEEE Global Communications Conference | 978-1-7281-8298-8/20/$31.00 ©2020 IEEE | DOI: 10.1109/GLOBECOM42002.2020.9348200

***Abstract*—Recent technological advances in the use of Unmanned Aerial Vehicles (UAVs) and Wireless Powered Communications (WPC) have enabled the energy efficient operation of the Public Safety Networks (PSN) during disaster scenarios. In this paper, an energy efficient information flow and energy harvesting framework capturing users' risk-aware characteristics is introduced based on the principles of Contract Theory. To better support the operational effectiveness of the proposed framework, users are clustered in rescue groups following a socio-physical-aware group formation mechanism, while rescue leaders for each group are selected. A reinforcement learning approach is applied to enable the optimal matching between the UAVs and the rescue leaders in a distributed and efficient manner. The proposed contract-theoretic framework models the UAVs-victims relation based on a labor market setting via offering rewards to the users (incentives) in order to compensate them for their invested labor (reporting information). Detailed numerical results demonstrate the benefits and superiority of the proposed framework under different settings.**

***Index Terms*—Contract Theory; Reinforcement Learning; Public Safety Systems; Resource Control**

## I. Introduction

Public Safety Networks (PSNs) have been introduced to provide reliable exchange of data during catastrophic events (e.g., natural disasters, terrorist attacks). The persistent and robust information flow in disaster-struck areas has been enabled by the usage of Unmanned Aerial Vehicles (UAVs). UAV-enabled wireless communications have attracted great research and commercial interest due to their salient attributes, i.e., controllable mobility, line-of-sight communication with the transmitters, and low-cost, fast, and flexible deployment [1]. Moreover, the Wireless Powered Communications (WPC) networking paradigm enables the mobile devices to harvest energy from the radio frequency signals of the transmitter [2]. Capitalizing on the advances achieved by these technologies, in this paper, we consider a UAV-assisted WPC network that enables the efficient data collection from a disaster-struck area, following a contract-theoretic approach.

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

### *A. Related Work & Motivation*

The problem of maximizing the system's energy efficiency in a three layer UAV-assisted network architecture (space-air-ground) is studied in [3] considering an Internet of Remote Things network, where the UAVs act as relays. The authors formulate and solve an optimization problem to determine the devices' subchannel selection, their optimal transmission power, and the UAVs' deployment. In [4], a UAV performs the data collection from an Internet of Things (IoT) field. The authors jointly optimize the UAV's flying speed, altitude, and the IoT devices' frame length at the MAC layer, to maximize the ground sensors energy efficiency. In [5], an ant colony optimization algorithm is presented that enables the collaboration between the UAVs and the ground devices, in order to prolong the lifetime of the network, by reducing the devices' energy consumption to report their data to the UAVs.

The concept of UAV-enabled WPC system has been introduced in [6], where UAV-mounted energy transmitters, transmit radio frequency signals and the ground devices harvest energy from them. In [7], the UAV's trajectory is obtained to maximize the harvested energy by the ground devices under the UAV's flying speed and altitude constraints. In [8], the authors aim at maximizing the minimum achievable throughput of the ground devices, by jointly optimizing the UAVs' trajectories, the users' transmission power, and the decision between the devices' energy harvesting and information transmission phases.

However, all these research efforts have been conducted in isolation focusing on only one of the following related problems, that is: the energy efficient information acquisition from the ground nodes, the energy harvesting from the UAVs' radio frequency signals, and/or the optimal UAVs' deployment. This fragmentation has not yet allowed the exploitation of the corresponding achievements in their full capacity. Accordingly, in this paper, we aim to address this research gap by introducing an energy efficient information flow and energy harvesting framework capturing users' risk-aware characteristics, based on the principles of *Contract Theory*, and the support of *Reinforcement Learning*.

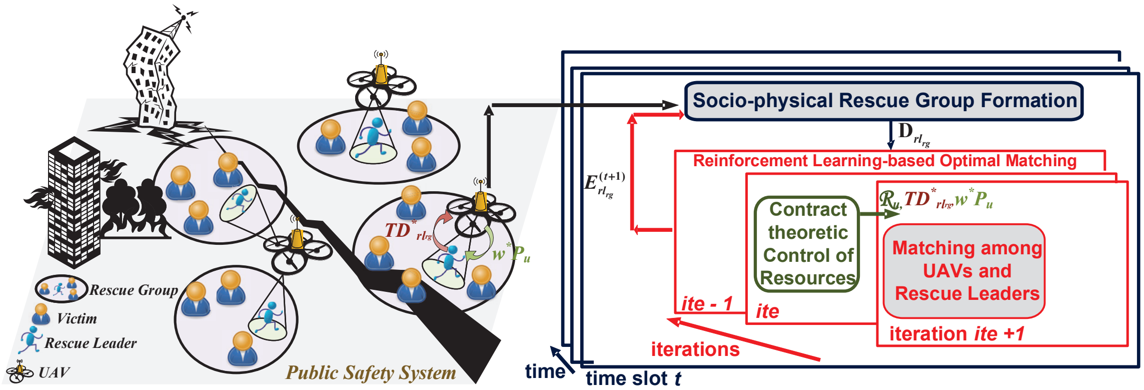


Fig. 1: UAV-assisted WPCN topology and framework's architecture

### B. Contributions & Outline

The main contributions of this research work are summarized as follows.

- A wireless powered communication network (WPCN) assisted by UAVs charging the victims' devices is considered in a public safety scenario. The victims' risk-aware characteristics to provide their information to the UAVs are captured in representative utility functions. An optimization problem determining each victim's optimal amount of provided information to the UAV and each UAV's optimal charging power, is formulated and solved following the principles of Contract Theory, by introducing a labor market relationship among the UAVs and the users (Section II).
- The victims are organized in rescue groups and the rescue leaders are determined for each group through a socio-physical groups formation mechanism (Section III-A). A reinforcement learning framework, based on the theory of *Stochastic Learning Automata*, is introduced to enable the optimal matching between the UAVs and the rescue leaders of each group, in a distributed and efficient manner (Section III-B).
- A set of simulation experiments are performed demonstrating the basic characteristics of the proposed contract-theoretic framework, while considering users' risk-aware behavior. The benefits of the proposed framework are highlighted in terms of energy-efficiency, information acquisition from the disaster area, and intelligent users' incentivization to support the rescue operation (Section IV).

## II. Contract-theoretic Control of Resources

A UAV-assisted WPCN is considered within a public safety system consisting of a set of victims $V = \{1, \dots, v, \dots, |V|\}$, a set of UAVs $U = \{1, \dots, u, \dots, |U|\}$, and the Emergency Control Center (ECC). The channel gain between two victims $v, v'$ is defined as $G_{v,v'} = \frac{\lambda}{d^2_{v,v'}}$, where $\lambda > 0$ represents the channel fading and $d_{v,v'}$ [m] is the distance among the victims $v$ and $v'$ [9]. Let $E_v$ [J] denote the energy availability of each victim's $v$ device and $d_v$ [m] represent the distance of the victim from the source of the disaster (e.g., epicenter of an earthquake). The victims are organized in rescue groups. Each rescue group $rg$ determines its rescue leader $rl_{rg}$ following a socio-physical rescue groups formation mechanism (Section III-A). Each rescue leader selects in a distributed manner to which UAV it will offload its data based on a reinforcement learning approach (Section III-B). The considered system's topology is presented in Fig. 1. Initially, we assume that the rescue groups formation and the rescue leaders association to the UAVs have already been performed and we focus on the contract-theoretic control of the resources.

During a catastrophic event, the ECC needs to collect information from the victims in order to plan the rescue operation [10]. Thus, incentives should be offered to them in order to provide information to the UAVs and correspondingly to the ECC. At the same time, the victims' behavioral characteristics, i.e., risk-aware behavior in terms of providing information, should be considered, while designing their incentives. To achieve this goal, the principles of Contract Theory are adopted [11]. Contract Theory is a powerful tool to design effective incentives by modeling the UAVs-victims relation based on a labor market setup. Specifically, the victims of a rescue group report their information to the corresponding rescue leader. Then, a UAV, which collects information from the rescue leader, considers the rescue leader's risk averse characteristics and offers rewards (i.e., incentives) in order to compensate it for its invested labor (i.e., reporting information).

Each victim transmits with power proportional to the normalized distance from its rescue leader, i.e., $P_v = \frac{d_{v,rl_{rg}}}{\max\limits_{v \in V_{rg}} d_{v,rl_{rg}}} \cdot P_v^{max}$, where $P_v^{max}$ is the victim's maximum transmission power and $V_{rg}$ is the set of victims belonging to the rescue group $rg$. The corresponding achievable transmission data rate is $R_v = W \cdot \log(1 + \frac{G_{v,rl_{rg}} P_v}{\sum\limits_{v' \geq v+1}^{|V'|} G_{v',rl_{rg}} P_{v'} + I_0})$, where $I_0$ represents the Additive White Gaussian Noise and $W$ [Hz] is the system's bandwidth, while non-orthogonal multiple access has been considered, and the successive interference cancellation technique is implemented at the receiver, i.e., rescue leader.

Thus, during a timeslot $t$ [sec], the total amount of data that the rescue leader collects is: $\mathbb{D}_{rl_{rg}} = (\sum_{v \in V_{rg}} R_v)t$ [bits].

Each UAV offers a contract to each rescue leader that is associated with. The contract is defined as $(w_{rl_{rg}} \cdot P_u^{max}, TD_{rl_{rg}})$, where $P_u^{max}$ is the UAV's maximum charging power and $TD_{rl_{rg}}$ are the collected data from its rescue group, where $TD_{rl_{rg}} \leq \mathbb{D}_{rl_{rg}}$. We consider the UAV's provided reward as $w_{rl_{rg}} \in [0, 1]$, thus, the corresponding charging power is $w_{rl_{rg}} \cdot P_u^{max}$. Each rescue leader invests an effort (i.e., labor) $a_{rl_{rg}} \in [0, 1]$ and transmits $TD_{rl_{rg}} = a_{rl_{rg}} \cdot \mathbb{D}_{rl_{rg}}$ data to the UAV. The rescue leader's performance, as it is evaluated by the UAV, is defined as $q_{rl_{rg}} = a_{rl_{rg}} + \epsilon$, where $\epsilon$ represents some noisy data. The parameter $\epsilon$ follows a normal distribution with zero mean and variance $\sigma^2$. Towards capturing the rescue leader's risk aware characteristics in terms of reporting information to the UAV, as well as its perceived satisfaction from its action and the harvested energy from the UAV, the rescue leader's risk aware utility function is defined as follows [11].

$$U_{rl_{rg}}(w_{rl_{rg}}, a_{rl_{rg}}) = -e^{-n_{rl_{rg}}[w_{rl_{rg}} - \psi(a_{rl_{rg}})]} \tag{1}$$

where $n_{rl_{rg}} \in (0, 1]$ is the rescue leader's risk aversion parameter. The greater the value of $n_{rl_{rg}}$ is, the more conservative the rescue leader becomes in terms of uploading information to the UAV in order to save its own energy. The function $\psi(a_{rl_{rg}})$ is the cost function of the rescue leader capturing its personal cost (energy consumption) to report the collected information from the disaster area to the UAV. The cost function is concave with respect to the rescue leader's invested effort, e.g., $\psi(a_{rl_{rg}}) = \frac{ca^2_{rl_{rg}}}{2}$, where $c > 0$ is a constant cost factor. The reward percentage $w_{rl_{rg}}$ offered by the UAV is defined as $w_{rl_{rg}} = \mu + s_{rl_{rg}} \cdot q_{rl_{rg}}$, where $\mu$ is a fixed compensation level, i.e., $\mu \cdot P_u^{max}$, to reward the rescue leaders for even participating in the information flow process, and $s_{rl_{rg}}$ is the variable compensation related to the rescue leader's performance component. The contract-theoretic control problem of the UAVs (i.e., charging power) and the rescue leaders (i.e., transmitted data) resources is formulated as a maximization problem of the UAV's expected profit.

$$\max_{a_{rl_{rg}}, s_{rl_{rg}}} \mathbb{E}(q_{rl_{rg}} - w_{rl_{rg}}) \tag{2a}$$

$$\mathbb{E}(-e^{-n_{rl_{rg}}[w_{rl_{rg}} - \psi(a_{rl_{rg}})]}) \geq U_{rl_{rg}}|_{min} \tag{2b}$$

$$a_{rl_{rg}} \in \arg\max_{a_{rl_{rg}}} \mathbb{E}(-e^{-n_{rl_{rg}}[w_{rl_{rg}} - \psi(a_{rl_{rg}})]}) \tag{2c}$$

where $U_{rl_{rg}}|_{min}$ is the minimum acceptable utility by the rescue leader in order to be motivated to send the collected data. The constraint (2b) represents the *individual rationality* constraint of the rescue leader. If this inequality does not hold true, then, the rescue leader has no incentive to report the collected data to the UAV. The constraint (2c) captures the *incentive compatibility* for each victim, i.e., each victim will put an effort to report the collected data in order to maximize its own perceived utility.

The rescue leader's expected utility can be written as $\mathbb{E}(-e^{-n_{rl_{rg}}[w_{rl_{rg}} - \psi(a_{rl_{rg}})]}) = -e^{-n_{rl_{rg}}[\mu + s_{rl_{rg}} a_{rl_{rg}} - \frac{ca^2_{rl_{rg}}}{2} - \frac{n_{rl_{rg}} s^2_{rl_{rg}} \sigma^2}{2}]}$ given that we can show that $\mathbb{E}(-e^{-n_{rl_{rg}} s_{rl_{rg}} \epsilon}) = e^{\frac{n_{rl_{rg}} s^2_{rl_{rg}} \sigma^2}{2}}$ from the theory of the normal distribution. Thus, by solving the constraint (2c), we can determine the rescue leader's optimal amount of transmitted data to the UAV.

$$TD^*_{rl_{rg}} = a^*_{rl_{rg}} \cdot \mathbb{D}_{rl_{rg}} = \frac{s_{rl_{rg}}}{c} \cdot \mathbb{D}_{rl_{rg}} \tag{3}$$

We can eliminate the constraint (2c) by substituting Eq. 3 to Eq. 2a and rewrite the optimization problem.

$$\max_{a_{rl_{rg}}, s_{rl_{rg}}} [\frac{s_{rl_{rg}}}{c} - (\mu + \frac{s^2_{rl_{rg}}}{c})] \tag{4a}$$

$$s.t. \quad \mu + \frac{s^2_{rl_{rg}}}{c} - \frac{c}{2}\frac{s^2_{rl_{rg}}}{c} - \frac{n_{rl_{rg}}}{2}\sigma^2 s^2_{rl_{rg}} = w_{rl_{rg}} \tag{4b}$$

The solution of the optimization problem (4a, 4b) yields to the optimal UAV's reward, i.e., charging power.

$$w^*_{rl_{rg}} \cdot P_u^{max} = [\mu + \frac{1}{1 + n_{rl_{rg}} c\sigma^2}(\frac{s_{rl_{rg}}}{c} + \epsilon)]P_u^{max} \tag{5}$$

Thus, the optimal contract among a UAV and rescue leader is $(w^*_{rl_{rg}} \cdot P_u^{max}, TD^*_{rl_{rg}})$. The operational timeslot of the system is splitted into the wireless energy transfer (WET) phase with duration $\tau_h$[sec] and the wireless information transmission (WIT) phase with duration $\tau_t$ [sec]. During the WET phase, the UAVs transfer directed energy to the rescue leaders that they are associated with, by unicasting a radio frequency signal via directional antennas [12]. The rescue leader's device's harvested energy from the UAV that it is associated with is given as follows.

$$HE_{rl_{rg}} = Eff_{rl_{rg}} \cdot \tau_h \cdot w_{rl_{rg}} * \cdot P_u^{max} \cdot G_{rl_{rg}, u} \tag{6}$$

where $Eff_{rl_{rg}} \in (0, 1]$ is the energy conversion efficiency factor, which depends on the rescue leader's device.

During the WIT phase, each rescue leader reports $TD^*_{rl_{rg}}$ to the UAV, assuming that its available energy, i.e., $E_{rl_{rg}} + HE_{rl_{rg}}$, is sufficient to report the contract theoretic optimal amount of data. Each rescue leader reports its optimal amount of data $TD^*_{rl_{rg}}$ through a dedicated subchannel with bandwidth $W$ [Hz] to the UAV via adopting the single carrier frequency division multiple access (SC-FDMA) technique. Thus, its available data rate is $W \cdot \log(1 + \frac{G_{rl_{rg}, UAV} P^{tr}_{rl_{rg}}}{I_0})$, where $P^{tr}_{rl_{rg}}$ is the rescue leader's transmission power. Thus, the rescue leader's consumed energy to transmit the $TD^*_{rl_{rg}}$ data is $E^{tr}_{rl_{rg}} = P^{tr}_{rl_{rg}} \cdot \tau_t$, and its remaining energy for the next timeslot is $E^{(t+1)}_{rl_{rg}} = E^{(t)}_{rl_{rg}} + HE_{rl_{rg}} - E^{tr}_{rl_{rg}}$.

## III. Groups Formation and UAV Associations

### A. Socio-physical Rescue Groups Formation

In this section, a socio-physical-aware rescue groups formation mechanism is presented, in order to enable the victims to create rescue groups and support the energy efficient information flow from the victims to the UAVs. In each rescue group, the victims transmit their information to the rescue leader of the group, who forwards it along with its own information to a UAV.

(1) Physical Ties: To support the victims' energy efficient communication, the victims tend to participate in rescue groups, where their communication distance among each other is small and their channel gain conditions are good. Thus, we define a symmetric matrix $G = \{g_{v,v'}\}_{|V| \times |V|}$, where $g_{v,v'} = \frac{G_{v,v'}}{\max\limits_{\forall v,v' \in V}\{G_{v,v'}\}} \in [0,1]$, which represents the normalized channel gain conditions of a pair of victims $v, v'$. Also, the victim's normalized energy availability $EA_v = \frac{E_v}{\max\limits_{\forall v' \in V}\{E_{v'}\}} \in [0,1]$ is critical in order to identify whether it could act as a rescue leader. The rescue leaders collect, process, and transmit the rest of the rescue group's victims' information, thus, they spend an increased amount of energy. Moreover, the victim's normalized distance from the source of the disaster $D_v = \frac{d_v}{\max\limits_{\forall v' \in V}\{d_{v'}\}} \in [0,1]$ is considered, as this victim can provide more accurate information to the UAV.

(2) Social Ties: The victims have interest to communicate with specific people, e.g., family members. The symmetric matrix $CI = \{ci_{v,v'}\}_{|V| \times |V|}, ci_{v,v'} \in [0,1]$ captures the victims' communication interest. A lower value of $ci_{v,v'}$ represents less communication interest among the victims.

By combining the victims' social and physical ties, we define a metric that captures the rescue and communication capability (RCC) of each victim, as follows.

$$RCC_v = EA_v \cdot D_v \cdot \sum_{v,v' \in V, v \neq v'} (ci_{v,v'} \cdot g_{v,v'}) \tag{7}$$

The socio-physical rescue groups formation mechanism is executed at the ECC, which informs the victims through broadcasted messages, and consists of the following steps.

(1) Initially, all the victims $|V|$ create a rescue group $rg$, whose set of victims is $V' = V$.

(2) For this rescue group $rg$ with set of victims $V'$, the rescue leader $rl_{rg}$ is determined as $rl_{rg} = \arg\max\limits_{v \in V'}\{RCC_v\}$.

(3) For the victims that belong to the rescue group $rg$ with set of victims $V'$, the following condition must hold true:

$$g_{v,rl_{rg}} \cdot ci_{v,rl_{rg}} \cdot D_v \geq RG^{(V')}_{thres} \tag{8}$$

where $RG^{(V')}_{thres} = \frac{\sum\limits_{v \in V'} g_{v,rl_{rg}}}{|V'|} \cdot \frac{\sum\limits_{v \in V'} ci_{v,rl_{rg}}}{|V'|} \cdot \frac{\sum\limits_{v \in V'} D_v}{|V'|}$ is a threshold value to create homogeneous rescue groups in terms of consisting of victims with close distance, good channel conditions, high communication interest among each other, as well as contributing valuable information due to their proximity to the source of the disaster. The victims, who do not satisfy the condition (8), they form a new rescue group, with set of victims $V'' \subseteq V'$.

(4) Set $V' = V' - V''$ and if $|V'| > 1$, return to step 2, otherwise stop.

### *B. Reinforcement Learning-enabled Matching among UAVs and Rescue Leaders*

In this section, a reinforcement learning-based framework is introduced to enable the optimal matching among the UAVs and the rescue leaders in a distributed and computationally efficient manner. Each leader acts as a stochastic learning automaton (SLA) making decisions of selecting a UAV to offload its data. Each UAV is characterized by a reputation, which depends on the physical and communication characteristics of the overall examined public safety system, and it is given as follows.

$$\mathcal{R}_u = \frac{\dfrac{\sum\limits_{rl_{rg} \in V_u} w^{*(ite-1)}_{rl_{rg}} P_u^{max} \sum\limits_{\forall rl_{rg}} d_{rl_{rg},u} \sum\limits_{rl_{rg} \in V_u} TD^{*(ite-1)}_{rl_{rg}} FT_u R_u}{\sum\limits_{\forall rl_{rg}} w^{*(ite-1)}_{rl_{rg}} P_u^{max} \sum\limits_{rl_{rg} \in V_u} d_{rl_{rg},u} \sum\limits_{\forall rl_{rg}} TD^{*(ite-1)}_{rl_{rg}} \sum\limits_{\forall u} R_u}}{|V_u^{(ite-1)}|^3} \tag{9}$$

where $FT_u \in (0,1)$ and $R_u$ are the normalized flying time and the communications coverage radius of the UAV $u$, respectively, and $V_u$ is the set of rescue leaders being served by the UAV $u$. The physical notion of Eq. 9 is that a rescue leader prefers to offload its data $TD^*_{rl_{rg}}$ to a UAV $u$ that (a) collectively charges with high transmission power the rescue leaders that are connected to it; (b) is in close proximity; (c) has a long flying time and large communications coverage radius; (d) it tends to collect large amount of data; and (e) is not overcongested by other rescue leaders trying to simultaneously offload their data.

The probability of a rescue leader selecting the same UAV $u$ to offload its data $TD^*_{rl_{rg}}$ in the next iteration of the SLA algorithm is given by Eq. 10a and the probability of selecting a different UAV is given by Eq. 10b [13].

$$Pr^{(ite+1)}_{rl_{rg},u} = Pr^{(ite)}_{rl_{rg},u} + b\mathcal{R}^{(ite)}_u (1 - Pr^{(ite)}_{rl_{rg},u}), u^{(ite+1)}_{rl_{rg}} = u^{(ite)}_{rl_{rg}} \tag{10a}$$

$$Pr^{(ite+1)}_{rl_{rg},u} = Pr^{(ite)}_{rl_{rg},u} - b\mathcal{R}^{(ite)}_u Pr^{(ite)}_{rl_{rg},u}, u^{(ite+1)}_{rl_{rg}} \neq u^{(ite)}_{rl_{rg}} \tag{10b}$$

where $u^{(ite)}_{rl_{rg}}$ is the selected UAV $u$ by the rescue leader $rl_{rg}$ in the iteration $ite$ of the SLA algorithm and $0 < b < 1$ is the learning parameter that controls how fast the rescue leaders learn their optimal UAV matching. It is noted that the UAVs' reputation values are broadcasted by them to the rescue leaders to enable the latter execute the SLA algorithm in a distributed manner and eliminate the signaling overhead. The SLA algorithm converges when $Pr^{(ite)}_{rl_{rg},u} \geq Pr_{thr}, \forall rl_{rg}$ where $Pr_{thr}$ is a threshold value, which for the evaluation purposes in this paper is $Pr_{thr} = 0.95$. Then, each rescue leader offloads its data $TD^*_{rl_{rg}}$ to the selected UAV, as shown in Fig. 1.

## IV. Numerical Results

A detailed numerical evaluation illustrates the performance of the proposed framework in terms of the: impact of socio-physical parameters (Section IV-A), contract-theoretic and behavioral-aware resource control (Section IV-B), and benefits of reinforcement learning to implement the optimal matching of the UAVs with the rescue leaders (Section IV-C). We consider $\tau_h = 0.985$ sec, $\tau_t = 0.015$

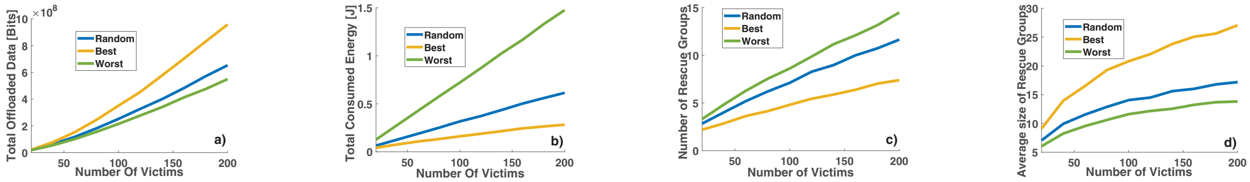


Fig. 2: Impact of socio-physical parameters under different comparative scenarios

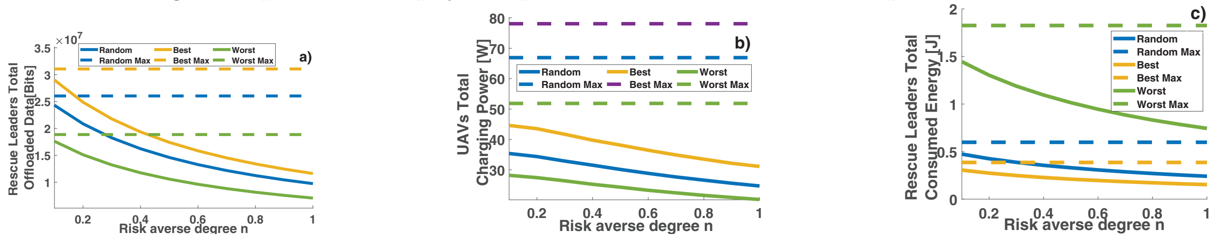


Fig. 3: (a) Rescue leaders' total amount of offloaded data, (b) UAVs' total charging power, and (c) Rescue leaders' total consumed energy w.r.t. risk averse degree for various comparative scenario

sec, $t = 1$ sec, $P_u^{max} = 85$W, $d_{v,v'} \in [30, 350]$m, $\lambda = 1$, $E_v \in [100, 400]$ J, $D_v \in [30, 350]$m, $W = 5 \cdot 10^6$Hz, $c = 1$, $b = 0.7$, $\mu = 0.5$, $FT_u \in (0, 1]$, and $R_u \in [30, 350]$m. We consider $|V| = 100$ victims, unless otherwise stated. The proposed framework's evaluation was conducted in a HP Laptop, 1.8GHz Intel Core i7, 16GB LPDDR3 RAM.

### A. *Impact of Socio-Physical Parameters*

Three comparative scenarios regarding the victims' socio-physical characteristics are evaluated: (i) *Best*: victims with high communication interest reside close to each other; (ii) *Worst*: victims with high communication interest reside far away from each other; and (iii) *Random*: victims have random communication interest and distance among each other. Fig. 2a-2d present the victims' data offloaded to their rescue leaders, their corresponding total consumed energy, the number of created rescue groups, and their corresponding average size, respectively, as a function of the number of victims for the three considered comparative scenarios. The results reveal that under the best case scenario, few homogeneous (in terms of the victims' socio-physical characteristics) rescue groups are created (Fig. 2c) of large average size (Fig. 2d), while the victims achieve to offload a large amount of data (Fig. 2a) with small consumed energy (Fig. 2b), due to their close proximity and good channel conditions among each other. The exact opposite holds true in the worst case scenario, while the random scenario presents an intermediate behavior between the best and worst case scenarios.

### B. *The Benefits of Contract Theory*

We also present the impact of the victims' risk-aware behavior on the resource management and the benefits of adopting contract theory to model the interactions among the UAVs and the rescue leaders. Six comparative scenarios are considered; three based on the proposed contract-theoretic resource control approach while assuming the best, worst, and random scenarios (Section IV-A), and the corresponding three scenarios that conclude by assuming that the UAVs charge the rescue leaders' devices with their maximum available power (referred to as Best Max, Worst Max, and Random Max respectively). Fig. 3a-3c present the rescue leaders' total amount of offloaded data, the UAVs' total charging power, and the rescue leaders' corresponding consumed energy to offload their data to the UAVs, respectively, as a function of the rescue leaders' risk averse degree, for all the considered comparative scenarios. It is observed that, with reference to the contract-theoretic based scenarios, as the rescue leaders become more risk averse (i.e., high value of the risk averse degree $n$), they tend to invest less effort in terms of offloading their data to the UAVs (Fig. 3a), thus, they consume less energy in their data transmission (Fig. 3c) and enjoy less rewards (i.e., charging power) from the UAVs (Fig. 3b). Also, in the comparative scenario, where the UAVs provide their maximum available charging power (Fig. 3b) to incentivize the rescue leaders to offload more data (Fig. 3a), this goal is achieved by immensely sacrificing the energy efficiency of the public safety system (PSS), as shown in Fig. 4a.

Specifically, Fig. 4a depicts the PSS's energy efficiency defined as the total amount of offloaded data by the rescue leaders over the corresponding spent charging power by the UAVs as a function of the rescue leaders' risk averse degree $n$. The results reveal that the UAVs' charging power is not well-spent, when they charge the rescue leaders with their maximum available charging power, and the UAVs' energy cost for every unit of collected information is higher for any examined topology of the PSS and the victims' socio-physical characteristics. This, demonstrates the benefit of the contract-theoretic control of the resources from the PSS's point of view. Moreover, the proposed framework is also valuable for the rescue leaders, as it enables them to achieve greater utility (Eq. 1) compared to the scenario of having their devices charged with the UAVs' maximum charging power (Fig. 4c).

### C. *Intelligent Matching between UAVs and Rescue Leaders*

In this subsection, we highlight the benefits of adopting a reinforcement learning mechanism to enable the rescue

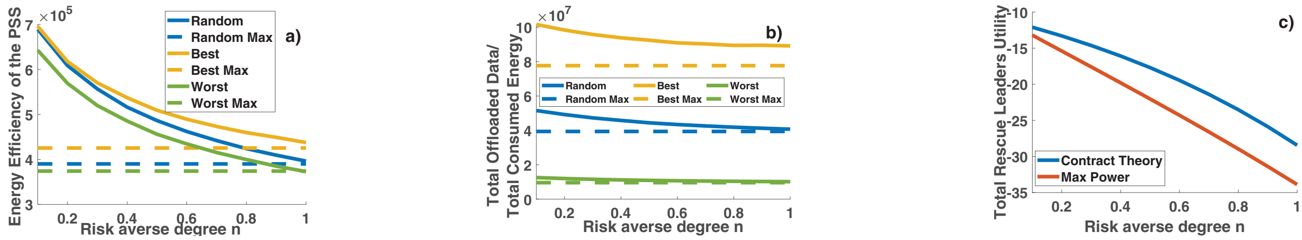


Fig. 4: (a) Energy Efficiency of the PSS, (b) Ratio of the rescue leaders total offloaded data over the total consumed energy, and (c) Total rescue leaders utility w.r.t. risk averse degree for various comparative scenario

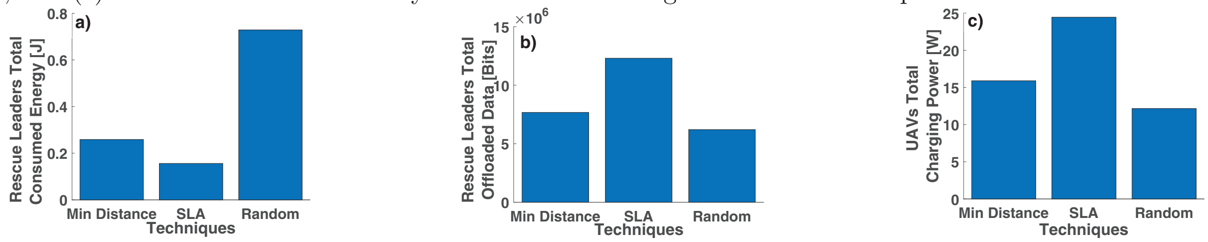


Fig. 5: Reinforcement learning-based matching between the UAVs and the rescue leaders – A comparative evaluation

leaders to optimally select a UAV to offload their data, while considering the characteristics of the PSS. Two indicative alternative approaches are also considered for comparison purposes: (a) *Min Distance*: the rescue leaders offload their data to the closest UAV; and (b) *Random*: the rescue leaders randomly select a UAV to offload their data. Fig. 5a-5c illustrate the rescue leaders' total consumed energy, their corresponding total amount of offloaded data, and the UAVs' total charging power, respectively, for the considered comparative scenarios. The results reveal that the reinforcement learning approach enables the rescue leaders to thoroughly learn their surrounding environment and make a sophisticated choice of a UAV, as indicated by the holistic reputation function (Eq. 9). Thus, the rescue leaders achieve to report a larger amount of data (Fig. 5b), compared to the other comparative scenarios, while consuming the lowest amount of energy (Fig. 5a), and enjoying greater charging power from the UAVs (Fig. 5c).

## V. Conclusions

In this paper, a resource orchestration framework is introduced in a UAV-assisted WPCN within a public safety system, based on the principles of contract theory and reinforcement learning. The key objective and novelty of this framework is that it enables the energy efficient information acquisition from the victims, while considering their risk-aware behavior. Detailed numerical results, obtained through modeling and simulation, demonstrate the benefits and superiority of the proposed framework in terms of energy-efficiency, information acquisition from the disaster area, and intelligent users' incentivization to support the rescue operation. Our future research plans include the extension of the proposed framework to consider the backhaul connection between the UAVs and the emergency control center, thus offering an energy efficient end-to-end data acquisition and transmission solution.

## References


[1] M. Mozaffari, W. Saad, M. Bennis, Y.-H. Nam, and M. Debbah, "A tutorial on uavs for wireless networks: Applications, challenges, and open problems," *IEEE communications surveys & tutorials*, vol. 21, no. 3, pp. 2334–2360, 2019.

[2] P.-V. Mekikis, A. Antonopoulos, E. Kartsakli, A. S. Lalos, L. Alonso, and C. Verikoukis, "Information exchange in randomly deployed dense wsns with wireless energy harvesting capabilities," *IEEE Transactions on Wireless Communications*, vol. 15, no. 4, pp. 3008–3018, 2016.

[3] Z. Li, Y. Wang, M. Liu, R. Sun, Y. Chen, J. Yuan, and J. Li, "Energy efficient resource allocation for uav-assisted space-air-ground internet of remote things networks," *IEEE Access*, vol. 7, pp. 145 348–145 362, 2019.

[4] X. Lin, G. Su, B. Chen, H. Wang, and M. Dai, "Striking a balance between system throughput and energy efficiency for uav-iot systems," *IEEE IoT Journal*, vol. 6, no. 6, pp. 10 519–10 533, 2019.

[5] K. Zhu, X. Xu, and Z. Huang, "Energy-efficient routing algorithms for uav-assisted mmtc networks," in *IEEE 30th PIMRC*. IEEE, 2019, pp. 1–6.

[6] J. Xu, Y. Zeng, and R. Zhang, "Uav-enabled wireless power transfer: Trajectory design and energy region characterization," in *2017 IEEE Globecom Workshops*. IEEE, 2017, pp. 1–7.

[7] ——, "Uav-enabled wireless power transfer: Trajectory design and energy optimization," *IEEE Transactions on Wireless Communications*, vol. 17, no. 8, pp. 5092–5106, 2018.

[8] J. Park, H. Lee, S. Eom, and I. Lee, "Uav-aided wireless powered communication networks: Trajectory optimization and resource allocation for minimum throughput maximization," *IEEE Access*, vol. 7, pp. 134 978–134 991, 2019.

[9] D. Sikeridis, E. E. Tsiropoulou, M. Devetsikiotis, and S. Papavassiliou, "Wireless powered public safety iot: A uav-assisted adaptive-learning approach towards energy efficiency," *Journal of Network and Computer Applic.*, vol. 123, pp. 69–79, 2018.

[10] G. Fragkos, E. E. Tsiropoulou, and S. Papavassiliou, "Disaster management and information transmission decision-making in public safety systems," in *IEEE GLOBECOM*. IEEE, 2019, pp. 1–6.

[11] P. Bolton, M. Dewatripont *et al.*, *Contract theory*. MIT press, 2005.

[12] Y. Wu, L. Qiu, and J. Xu, "Uav-enabled wireless power transfer with directional antenna: A two-user case," in *15th ISWCS*. IEEE, 2018, pp. 1–6.

[13] G. Fragkos, P. A. Apostolopoulos, and E. E. Tsiropoulou, "Escape: Evacuation strategy through clustering and autonomous operation in public safety systems," *Future Internet*, vol. 11, no. 1, p. 20, 2019.