2023 19th International Conference on Distributed Computing in Smart Systems and the Internet of Things (DCOSS-IoT)

2023 19th International Conference on Distributed Computing in Smart Systems and the Internet of Things (DCOSS-IoT) | 979-8-3503-4649-7/23/$31.00 ©2023 IEEE | DOI: 10.1109/DCOSS-IOT58021.2023.00105

Positioning, Navigation, and Timing on the Air

Md Sadman Siraj*, Aisha B Rahman*, Panagiotis Charatsaris†, Eirini Eleni Tsiropoulou*, Symeon Papavassiliou†
* Dept. of Electrical and Computer Engineering, University of New Mexico, Albuquerque, NM, USA
† School of Electrical and Computer Engineering, National Technical University of Athens, Athens, Greece

***Abstract*—The widespread adoption of Positioning, Navigation, and Timing (PNT) services in many applications across modern society, has given rise to the development of various alternative PNT systems, aiming at addressing the Global Positioning System (GPS) vulnerabilities. In this paper, we introduce a collaborative-based PNT solution by jointly exploiting the use of UAV-assisted wireless networks and the collaboration among the targets and the collaborator nodes in the network, who have unknown and a rough estimate of their position and timing, respectively. The minimization problem of the overall system's and each collaborator node's and target's position and timing estimation error is formulated and solved based on the principles of potential games. The existence of a Nash Equilibrium (NE) of the corresponding game is proven, while two different approaches to obtain the game's NE are studied and their tradeoffs are evaluated. The first one follows a game-theoretic approach and the second one is based on the principles of reinforcement learning. A detailed numerical evaluation is performed, via modeling and simulation, in order to demonstrate the benefits and tradeoffs of the proposed PNT solution.**

***Index Terms*—Positioning, Navigation, Timing, Unmanned Aerial Vehicles, Game Theory, Reinforcement Learning.**

I. Introduction

Positioning, Navigation, and Timing (PNT) is an important research topic, as the PNT services are necessary for almost every modern application, such as disaster management and rescue planning, smart transportation, healthcare monitoring, logistics and supply chain management, just to name few. Nowadays, the dominant PNT service provider is the Global Navigation Satellite System (GNSS), with the most representative example being the Global Positioning System (GPS) [1]. However, the applicability and accuracy of the GPS can be hampered either by man-made or unintentional interference to the satellite signals received by the targets, high power attenuation of the signals due to long propagation distances, spoofing, and jamming [2]. In such cases, GPS-denial phenomena are observed or the provided GPS-based PNT services' accuracy is very low. Thus, the need for development of alternative PNT solutions, which can seamlessly substitute or complement the GPS-based PNT services, is more pressing than ever. In this paper, a novel PNT on the air framework is introduced by exploiting the use of Unmanned Aerial Vehicles (UAVs), which act as flying anchor nodes. A collaborative-based PNT solution is designed enabling the different types of nodes in the system, to collaborate among each other in order to minimize their experienced positioning and timing error. Those nodes refer to the targets and collaborator nodes, who have unknown and a rough estimate of their positioning and timing, respectively. In particular, the UAVs and the collaborator nodes support the targets to improve the accuracy in determining their position and timing. The aforementioned optimization goal is treated through a game theoretic formulation, while two different approaches are studied in order to obtain the game's stable operation point. The first one follows a game-theoretic approach and the second one is based on the principles of reinforcement learning. The drawbacks and benefits of each one of the aforementioned techniques are also analyzed and evaluated.

This work was supported in part by the National Technical University of Athens Research Committee Grant on "Network Management and Optimization" (Award #95028000).
The research of Dr. Tsiropoulou was partially supported by the NSF Award #2219617.

A. Related Work

The design of alternative PNT solutions and systems has recently attracted the interests of the research community aiming at supporting both indoor and outdoor use case scenarios [3]. A fingerprint-based localization technique for indoor environments is introduced in [4] dealing with the problem of fingerprint similarity based on a deep neural network approach, where the original fingerprints are replaced by the hidden layer parameters. Similarly, in [5], a broad learning system is designed exploiting the targets' received signals' channel state information to deal with the problems of data loss and noise interference in the fingerprint database. The authors in [6] exploit the Bluetooth Low Energy (BLE) Received Signal Strength Indicator (RSSI) based on the targets' wearable wrist watch signals in order to feed a self-supervised machine learning model that performs the targets' localization. An indoor localization model is proposed in [7] using the concurrent angle of arrival estimation based on signals transmitted by ultra-wide band radios aiming at reducing the number of required packet exchanges in order to determine the targets' position.

Several other techniques have been introduced in the literature to design alternative PNT solutions by exploiting different types of communication technologies [8]. The Long-Range (LoRa) wireless technology is investigated in [9] to design an RSSI-based localization algorithm aiming at reducing the effect of Gaussian and non-Gaussian noise during the local-

2325-2944/23/$31.00 ©2023 IEEE
DOI 10.1109/DCOSS-IoT58021.2023.00105

ization process. The PNT services are supported by a robot in [10], which moves on a predefined trajectory, collects RSSI data from the targets, and determines their positions based on the signals' angle of arrival. Accurate PNT services are also very critical in maritime and underwater rescue operations [11]. The authors in [12] exploit the linear frequency modulated signals transmitted by underwater moving targets within a mobile underwater accoustic array network and they determine the targets' position velocity by exploiting the signals' propagation delay and Doppler effect. Similarly, in [13], the underwater targets localize themselves by accoustically polling beacon signals statically deployed at well-known locations following an Extended Kalman Filter approach. A channel charting-aided localization mechanism is introduced in [14] for millimeter wave networks by exploiting the multipath channel state information of the targets' received signals from at least four base stations.

Precise and real-time PNT services for vehicles and in general, outdoor moving targets, becomes challenging due to various factors, such as shadow areas, e.g., tunnels, high-rise buildings, densely vegetated areas, etc [15]. The authors in [16] deal with the problem of missing measurements of signals between multiple tags attached to a vehicle and anchor nodes by introducing a Euclidean distance matrix completion approach that determines the bounds of the missing measurements. An asynchronous advantage actor-critic algorithm is proposed in [17] based on the principles of reinforcement learning in order to determine the optimal targets' positioning by performing corrections on the raw GPS observations. Also, a particle swarm optimization technique is introduced in [18] to perform the localization of the Unmanned Aerial Vehicles (UAVs) by reducing the time complexity and localization error. A different approach is followed in [19], where the authors use drone aerial images to train a deep learning model in order to perform human subjects localization.

Recently, the concept of cooperation among targets or in general, localization-related equipment has been introduced as a novel approach to reduce the position and timing error [20]. In [21], a multi-hypothesis Extended Kalman Filter technique is proposed to enable the estimation of the relative position and orientation between vehicles, in cases of high initial uncertainty. In [22], a fully connected wireless powered communication network (WPCN) is considered, where the targets determine their positions by exploiting the signals transmitted by remote energy access points in order for all the targets to harvest energy. A cooperative PNT model among UAVs is discussed in [23], where a UAV, with known position, broadcasts periodically its position, in order for the other UAVs, with unknown position, to measure the signals' direction of arrival and ultimately, determine their positions. A low-cost high-performance distributed spatio-temporal information based on cooperative positioning algorithm is proposed in [24] for 3D wireless networks, supporting any type of ranging measurements, e.g., RSSI, angle of arrival, that can determine the relative position among targets.

B. Contributions & Outline

Though the current state-of-the-art has demonstrated tremendous progress in designing alternative PNT solutions, the joint problem of determining the targets' position and timing, i.e., clock difference among the targets and anchor nodes, remains highly unexplored. Also, the vast majority of the proposed alternative PNT solutions suffer from high computational complexity, e.g., deep learning-based models, or require dedicated equipment.

Towards addressing those issues, a novel PNT on the air framework is proposed in this paper by exploiting the UAV-assisted wireless networks and introducing collaboration among the targets and the collaborator nodes, who have unknown and a rough estimate of their position and timing, respectively. The ultimate goal of the PNT on the air framework is to accurately determine the position and timing of both the targets and collaborator nodes under both static and mobile use case scenarios. The novel key contributions of our research work are summarized below.

1) A novel collaborative-based PNT architecture is introduced consisting of UAVs, collaborator nodes and targets. The UAVs act as flying base stations, supporting the PNT services, while the collaborator nodes have a rough estimate of their position and timing. The UAVs and the collaborator nodes support the targets to accurately determine their position and timing, which initially are completely unknown to them.
2) A minimization problem of the overall system's and each collaborator node's and target's position and timing estimation error is formulated and solved based on the principles of potential games. Specifically, the problem is addressed as a non-cooperative game among the collaborator nodes and the targets, who select their accurate position and timing strategy in order to minimize the estimation error. The existence of a Nash Equilibrium (NE) for the non-cooperative game is proven.
3) A Synchronous (BRD) and Asynchronous (BRD) Best Response Dynamics algorithm are introduced to determine the NE. Also, following the principles of log-linear reinforcement learning (RL), two alternative RL algorithms are introduced to determine the NE based on the exploration and learning phases.
4) A detailed evaluation is performed, via modeling and simulation, to quantify the accuracy and time complexity of the proposed PNT on the air framework. A thorough study of the trade-off between time complexity and accuracy for the alternative introduced algorithmic approaches, i.e., Best Response Dynamics and RL, is provided, as well as a comparative evaluation of the proposed PNT on the air framework to the state-of-the-art is presented. Finally, the accuracy of the proposed model under targets' mobility scenario is evaluated.

The remainder of this paper is organized as follows, Section II presents the PNT on the air system model and provides an overview of the overall proposed framework. The minimiza-

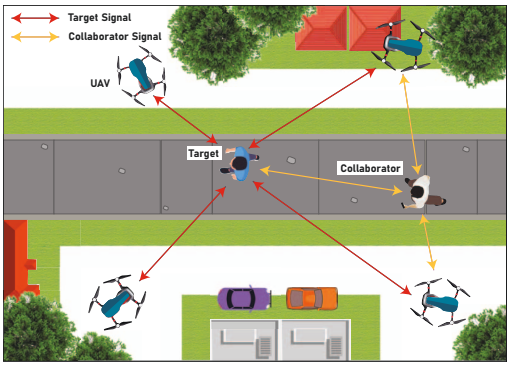

Fig. 1: Overview of the topology.

tion problem of the targets' and collaborator nodes' position and timing estimation error is formulated and solved in Section III, while different algorithmic approaches (i.e.e, BRD and/or RL based) that determine the NE are discussed in Section III-A and III-B, respectively. The detailed numerical evaluation is provided in Section IV, and Section V concludes the paper.

II. System Model

A novel PNT on the air system model is considered, consisting of a set of UAVs $\mathcal{U} = \{1, \ldots, u, \ldots, U\}$, a set of collaborator nodes (or simply called collaborators) $\mathcal{C} = \{1, \ldots, c, \ldots, C\}$, and a set of targets $\mathcal{T} = \{1, \ldots, t, \ldots, T\}$. An overview of the considered topology is presented in Fig. 1 [25]. The UAVs have perfect knowledge of their position and timing (PT) $\mathbf{P}_u = \{x_u, y_u, z_u, \Delta t_u\}$, while the collaborators have a rough estimate of their PT $\hat{\mathbf{P}}_c = \{\hat{x}_c, \hat{y}_c, \hat{z}_c, \hat{\Delta t}_c\}$. The targets have fully unknown PT $\hat{\mathbf{P}}_t = \{\hat{x}_t, \hat{y}_t, \hat{z}_t, \hat{\Delta t}_t\}$. The collaborators and the targets collaborate with each other, by exploiting the signals transmitted by the UAVs in order to accurately determine their PT [26]. The operation of the PNT on the air framework is demonstrated in Fig. 2.

Specifically, each target and collaborator $j \in \mathcal{T} \cup \mathcal{C}$, send a ranging request beacon signal with fixed transmission power P [W], which is received by the UAVs $\mathcal{U}_j \subseteq \mathcal{U}$ and other collaborators $\mathcal{C}_j \subseteq \mathcal{C}$ within the corresponding coverage area. Then, the corresponding UAVs $\mathcal{U}_j$ and collaborators $\mathcal{C}_j$ send a ranging reply beacon signal with transmission power P [W] including digital information of their PT, i.e., $\mathbf{P}_u, \forall u \in \mathcal{U}_j$, and $\hat{\mathbf{P}}_c, \forall c \in \mathcal{C}_j$. Through this process, each target and collaborator $j \in \mathcal{T} \cup \mathcal{C}$ has identified its neighborhood and has been informed about the PT of its neighbors, i.e., $\mathbf{P}_u, \forall u \in \mathcal{U}_j$, and $\hat{\mathbf{P}}_c, \forall c \in \mathcal{C}_j$. Then, each target and collaborator $j \in \mathcal{T} \cup \mathcal{C}$, can measure the pseudoranges $d_{j,c}, \forall c \in \mathcal{C}_j$ and $d_{j,u}, \forall u \in \mathcal{U}_j$ based on the received power:

$$P_{j,k}^{rec} = P \frac{G_k^{trans} G_j^{rec}}{L_{j,k}} \tag{1}$$

where $k = \{c, u\}, \forall c \in \mathcal{C}_j, \forall u \in \mathcal{U}_j$, G_k^{trans} denotes the gain of the transmitting node's k antenna, G_j^{rec} is the gain of the receiving node's j antenna [27], and $L_{j,k}$ captures the power attenuation model, which is given in Eq. 2, based on the Okumura — Hata model [28] for large cities scenarios, where f_c [Hz] is the carrier frequency, h_k^{trans} [m] is the height of the transmitting node k, h_j^{rec} [m] is the height of the receiving node, and $d_{j,k}$ [m] denotes the measured pseudorange, i.e., distance, among the target or collaborator j and the transmitting node k.

Based on this process, each target and collaborator knows the pseudoranges $d_{j,k}, \forall k \in \mathcal{U}_j \cup \mathcal{C}_j$, and the PT of each neighbor node, i.e., $\mathbf{P}_u, \forall u \in \mathcal{U}_j$, and $\hat{\mathbf{P}}_c, \forall c \in \mathcal{C}_j$. Then, by implementing the Iterative Least Square (ILS) algorithm [29], each target and collaborator can determine an initial estimate of their PT, i.e., $\hat{\mathbf{P}}_t, \forall t \in \mathcal{T}$, and $\hat{\mathbf{P}}_c, \forall c \in \mathcal{C}_j$.

III. Positioning, Navigation, and Timing on the Air

The goal of the proposed PNT on the air framework is to accurately determine the position and timing of all the targets and collaborators by minimizing the estimation error that is experienced by each one of them, and by jointly minimizing the estimation error in the overall examined system. Initially, we define the Euclidean distance of the PT estimations among all the involved entities in the PNT system, as follows:

$$\hat{d}(\hat{\mathbf{P}}_j, \hat{\mathbf{P}}_k) = \begin{cases} ||\hat{\mathbf{P}}_j - \mathbf{P}_k||, \text{ if } k \in \mathcal{U}_j \\ ||\hat{\mathbf{P}}_j - \hat{\mathbf{P}}_k||, \text{ if } k \in \mathcal{C}_j \end{cases} \tag{3}$$

The Euclidean distance $\hat{d}(\hat{\mathbf{P}}_j, \hat{\mathbf{P}}_k)$ is an estimation performed by each target and collaborator $j \in \mathcal{T} \cup \mathcal{C}$, based on the neighborhood identification process, as presented in Section II. Also, each target and collaborator has measured the distance from its neighbors and can receive the clock offset information from them through the neighborhood identification process, thus, the corresponding measured distance $d_{j,k}, \forall j \in \mathcal{T} \cup \mathcal{C}, \forall k \in \mathcal{U}_j \cup \mathcal{C}_j$. Therefore, the position and timing estimation error experienced by each target and collaborator j, is derived as follows:

$$\epsilon(\hat{\mathbf{P}}_j, \hat{\mathbf{P}}_k) = [d_{j,k} - \hat{d}(\hat{\mathbf{P}}_j, \hat{\mathbf{P}}_k)]^2 \tag{4}$$

The goal of each target and collaborator $j \in \mathcal{T} \cup \mathcal{C}$ is to minimize its experienced position and timing estimation error, thus, the corresponding optimization problem is formulated as follows:

$$\min_{\{\hat{\mathbf{P}}_j\}_{\forall j \in \mathcal{T} \cup \mathcal{C}}} \sum_{\forall k \in \mathcal{U}_j \cup \mathcal{C}_j} \epsilon(\hat{\mathbf{P}}_j, \hat{\mathbf{P}}_k) \tag{5}$$

From a system's perspective, the goal of the system is to minimize the overall system's position and timing estimation error, thus, the corresponding optimization problem is formulated as follows:

$$\min_{\{\hat{\mathbf{P}}_j\}_{\forall j \in \mathcal{T} \cup \mathcal{C}}} F(\hat{\mathbf{P}}_j, \hat{\mathbf{P}}_k) = \sum_{\forall j \in \mathcal{T} \cup \mathcal{C}} \sum_{\forall k \in \mathcal{U}_j \cup \mathcal{C}_j} \epsilon(\hat{\mathbf{P}}_j, \hat{\mathbf{P}}_k) \tag{6}$$

Towards solving the optimization problems presented in Eq. 5 and Eq. 6, we formulate the interactions among the targets and the collaborators as a non-cooperative game $G = [\mathcal{J}, \{\mathcal{S}_j\}_{\forall j \in \mathcal{J}}, \{U_j\}_{\forall j \in \mathcal{J}}]$, where $\mathcal{J} = \mathcal{T} \cup \mathcal{C}$ is the set of

$$L_{j,k} = 69.55 + 26.16 \log f_c + (44.9 - 6.55 \log h_k{}^{trans}) \log d_{j,k} - 13.82 \log h_k^{trans} - 3.2[log(11.75 h_j^{rec})]^2 - 4.97[dB] \quad (2)$$

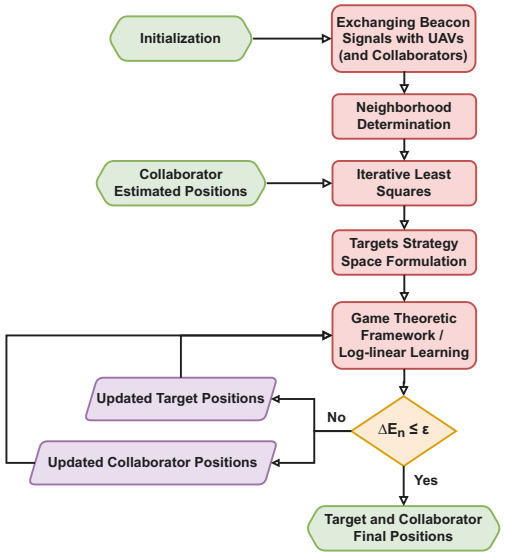

Fig. 2: PNT on the air framework.

players, i.e., targets and collaborators, $\mathcal{S}_j$ is the strategy set of positioning and timing strategies for node j, and $U_j(\mathbf{s}_j, \mathbf{s}_{-j})$ denotes the node's j payoff function, where $U_j(\mathbf{s}_j, \mathbf{s}_{-j}) = \sum_{\forall k \in \mathcal{U}_j \cup \mathcal{C}_j} \epsilon(\hat{\mathbf{P}}_j, \hat{\mathbf{P}}_k)$.

Definition 1. ***(Nash Equilibrium)*** *A strategy vector* $\mathbf{s}^* = (\mathbf{s}_1^*, \ldots, \mathbf{s}_j^*, \ldots, \mathbf{s}_J^*)$ *is a Nash Equilibrium for the non-cooperative game* $G = [\mathcal{J}, \{\mathcal{S}_j\}_{\forall j \in \mathcal{J}}, \{U_j\}_{\forall j \in \mathcal{J}}]$, *iff* $U_j(\mathbf{s}_j^*, \mathbf{s}_{-j}^*) \geqslant U_j(\mathbf{s}_j^{'}, \mathbf{s}_{-j}^*)$, $\forall \mathbf{s}_j^{'} \in \mathcal{S}_j$, *where* $\mathbf{s}_{-j}^* = [\mathbf{s}_1^*, \ldots, \mathbf{s}_{j-1}^*, \mathbf{s}_{j+1}^*, \ldots, \mathbf{s}_J^*]$.

Towards showing the existence of at least one Nash Equilibrium for the non-cooperative game $G = [\mathcal{J}, \{\mathcal{S}_j\}_{\forall j \in \mathcal{J}}, \{U_j\}_{\forall j \in \mathcal{J}}]$, we use the theory of potential games.

Definition 2. ***(Exact Potential Game)*** *A non-cooperative game* G *is an exact potential game iff:*

$$\Phi(\mathbf{s}_j, \mathbf{s}_{-j}) - \Phi(\mathbf{s}_j^{'}, \mathbf{s}_{-j}) = U_j(\mathbf{s}_j, \mathbf{s}_{-j}) - U_j(\mathbf{s}_j^{'}, \mathbf{s}_{-j})$$

where $\Phi(\mathbf{s}_j, \mathbf{s}_{-j})$ *is the potential function.*

Focusing on the non-cooperative game $G = [\mathcal{J}, \{\mathcal{S}_j\}_{\forall j \in \mathcal{J}}, \{U_j\}_{\forall j \in \mathcal{J}}]$, the following theorem proves that it is an exact potential game, thus, it admits at least one Nash Equilibrium [30].

Theorem 1. *The non-cooperative game* $G = [\mathcal{J}, \{\mathcal{S}_j\}_{\forall j \in \mathcal{J}}, \{U_j\}_{\forall j \in \mathcal{J}}]$ *is an exact potential game with potential function:*

$$\Phi(\mathbf{s}_j, \mathbf{s}_{-j}) = \frac{F(\mathbf{s}_j, \mathbf{s}_{-j})}{2} = \frac{\sum\limits_{\forall j \in \mathcal{T} \cup \mathcal{C}} \sum\limits_{\forall k \in \mathcal{U}_j \cup \mathcal{C}_j} \epsilon(\mathbf{s}_j, \mathbf{s}_k)}{2}.$$

Proof: Initially, we determine the difference of node's j payoff function for two alternative strategies $\mathbf{s}_j, \mathbf{s}_j^{'}$, while the rest of the nodes keep their strategies the same, i.e., $\mathbf{s}_{-j}$.

$$U_j(\mathbf{s}_j, \mathbf{s}_{-j}) - U_j(\mathbf{s}_j^{'}, \mathbf{s}_{-j}) = \sum_{\forall k \in \mathcal{N}_j} \epsilon(\mathbf{s}_j, \mathbf{s}_k) - \sum_{\forall k \in \mathcal{N}_j} \epsilon(\mathbf{s}_j^{'}, \mathbf{s}_k)$$

where $\mathcal{N}_j = \mathcal{U}_j \cup \mathcal{C}_j$ and $\mathcal{J} = \mathcal{T} \cup \mathcal{C}$. Then, we analyze the potential function, as follows [31]:

$$\Phi(\mathbf{s}_j, \mathbf{s}_{-j}) = \frac{1}{2} \sum_{\forall j \in \mathcal{J}} \sum_{\forall k \in \mathcal{N}_j} \epsilon(\mathbf{s}_j, \mathbf{s}_k) = \frac{1}{2}[\sum_{\forall k \in \mathcal{N}_j} \epsilon(\mathbf{s}_j, \mathbf{s}_k) + \sum_{\substack{\forall n \in \mathcal{J} \\ n \neq j}} \sum_{\forall k \in \mathcal{N}_j} \epsilon(\mathbf{s}_n, \mathbf{s}_k)] = \frac{1}{2}[\sum_{\forall k \in \mathcal{N}_j} \epsilon(\mathbf{s}_j, \mathbf{s}_k) + \sum_{\substack{\forall n \in \mathcal{J} \\ n \neq j}} [(\sum_{\substack{\forall k \in \mathcal{N}_n \\ k \neq j}} \epsilon(\mathbf{s}_n, \mathbf{s}_k)) + \epsilon(\mathbf{s}_n, \mathbf{s}_j)]] = \frac{1}{2}[\sum_{k \in \mathcal{N}_j} \epsilon(\mathbf{s}_j, \mathbf{s}_k) + \sum_{\substack{\forall n \in \mathcal{J} \\ n \neq j}} \sum_{\substack{\forall k \in \mathcal{N}_n \\ k \neq j}} \epsilon(\mathbf{s}_n, \mathbf{s}_k) + \sum_{\substack{\forall n \in \mathcal{J} \\ n \neq j}} \epsilon(\mathbf{s}_n, \mathbf{s}_j)].$$

It is noted that if two nodes n, j are not neighbors, then, they cannot measure the distances among each other, thus:

$$\epsilon(\mathbf{s}_n, \mathbf{s}_j) = 0, \text{ if } n, j \notin \mathcal{N}_j$$

Based on this observation, the last term of the potential function can be written as follows:

$$\sum_{\substack{\forall n \in \mathcal{J} \\ n \neq j}} \epsilon(\mathbf{s}_n, \mathbf{s}_j) = \sum_{\forall n \in \mathcal{N}_j} \epsilon(\mathbf{s}_n, \mathbf{s}_j) + \underbrace{\sum_{\substack{\forall n \notin \mathcal{N}_j \\ n \neq j}} \epsilon(\mathbf{s}_n, \mathbf{s}_j)}_{=0} = \sum_{\forall n \in \mathcal{N}_j} \epsilon(\mathbf{s}_n, \mathbf{s}_j).$$

Therefore, the potential function can be written as follows:

$$\Phi(\mathbf{s}_j, \mathbf{s}_{-j}) = \frac{1}{2}[\sum_{\forall k \in \mathcal{N}_j} \epsilon(\mathbf{s}_j, \mathbf{s}_k) + \sum_{\substack{\forall n \in \mathcal{J} \\ n \neq j}} \sum_{\substack{\forall k \in \mathcal{J}_n \\ k \neq j}} \epsilon(\mathbf{s}_n, \mathbf{s}_k) + \sum_{\forall n \in \mathcal{N}_j} \epsilon(\mathbf{s}_n, \mathbf{s}_j)] = \frac{1}{2}[2 \sum_{\forall k \in \mathcal{N}_j} \epsilon(\mathbf{s}_j, \mathbf{s}_k) + \sum_{\substack{\forall n \in \mathcal{J} \\ n \neq j}} \sum_{\substack{\forall k \in \mathcal{N}_n \\ k \neq j}} \epsilon(\mathbf{s}_n, \mathbf{s}_k)] = \sum_{\forall k \in \mathcal{N}_j} \epsilon(\mathbf{s}_j, \mathbf{s}_k) + \frac{1}{2} \sum_{\substack{\forall n \in \mathcal{J} \\ n \neq j}} \sum_{\substack{\forall k \in \mathcal{N}_n \\ k \neq j}} \epsilon(\mathbf{s}_n, \mathbf{s}_k).$$

By taking the difference of the potential function for two alternative strategies $\mathbf{s}_j, \mathbf{s}_j^{'}$, we have:

$$\Phi(\mathbf{s}_j, \mathbf{s}_{-j}) - \Phi(\mathbf{s}_j^{'}, \mathbf{s}_{-j}) = \sum_{\forall k \in \mathcal{N}_j} \epsilon(\mathbf{s}_j, \mathbf{s}_k) + \frac{1}{2} \sum_{\substack{\forall n \in \mathcal{J} \\ n \neq j}} \sum_{\substack{\forall k \in \mathcal{N}_n \\ k \neq j}} \epsilon(\mathbf{s}_n, \mathbf{s}_k) - \sum_{\forall k \in \mathcal{N}_j} \epsilon(\mathbf{s}_j^{'}, \mathbf{s}_k) - \frac{1}{2} \sum_{\substack{\forall n \in \mathcal{J} \\ n \neq j}} \sum_{\substack{\forall k \in \mathcal{N}_n \\ k \neq j}} \epsilon(\mathbf{s}_n, \mathbf{s}_k) = \sum_{\forall k \in \mathcal{N}_j} \epsilon(\mathbf{s}_j, \mathbf{s}_k) - \sum_{\forall k \in \mathcal{N}_j} \epsilon(\mathbf{s}_j^{'}, \mathbf{s}_k) = U_j(\mathbf{s}_j, \mathbf{s}_{-j}) - U_j(\mathbf{s}_j^{'}, \mathbf{s}_{-j}).$$

Thus, the non-cooperative game $G = [\mathcal{J}, \{\mathcal{S}_j\}_{\forall j \in \mathcal{J}}, \{U_j\}_{\forall j \in \mathcal{J}}]$ is an exact potential game and has at least one Nash Equilibrium. ■

Algorithm 1 Synchronous Best Response Dynamics (SBRD)

Input: $\mathbf{P}_u$, $\forall u \in \mathcal{U}$, $\hat{\mathbf{P}}_c$, $\forall c \in \mathcal{C}$
Output: $\boldsymbol{s}^*$
Initialization: $i = 0$, $Convergence = 0$, $\mathbf{s}^{i=0}$ randomly selected strategy.
while *Convergence* $== 0$ **do**
  $i = i + 1$;
  for all $j \in \mathcal{J} = \mathcal{T} \cup \mathcal{C}$ **do**
    Determine $\mathbf{s}_j^{*i}$ *(Eq. 5) and* $U_j(\mathbf{s}_j^{*i}, \mathbf{s}_{-j}^{i-1})$ *(Eq. 4), given* $\mathbf{s_{-j}}^{i-1}$
  end for
  if $|U_j(\mathbf{s}_j^{*i}, \mathbf{s}_{-j}^{i-1}) - U_j(\mathbf{s}_j^{*i+1}, \mathbf{s}_{-j}^{i})| \leqslant \delta$, δ small positive number, $\forall j \in \mathcal{J}$ **then**
    $Convergence = 1$
  end if
end while

A. Game Theory enabling PNT

Aiming at determining the Nash Equilibrium, a Best Response Dynamics (BRD) approach can be followed, where each target and collaborator performs a best response in terms of selecting its position and timing, based on the strategies selected by the other nodes in the system, aiming at minimizing its experienced estimation error. If all the targets and collaborators perform simultaneously their best responses, then, the Synchronous BRD (SBRD) algorithm is implemented, as described in Algorithm 1. If the nodes perform their best responses in a sequential manner, then, the Asynchronous BRD (ABRD) is adopted, as presented in Algorithm 2. Those different decision-making patterns result in the herding effect experienced among the targets and the collaborators in the SBRD algorithm, resulting in worse estimation error, but faster convergence time (given that all the nodes update simultaneously their strategies) versus the ABRD algorithm. Detailed comparative evaluation of the drawbacks and benefits of the ABRD and SBRD algorithms are provided in Section IV.

B. A Reinforcement Learning-based Perspective

The Nash equilibrium of the non-cooperative game $G = [\mathcal{J}, \{\mathcal{S}_j\}_{\forall j \in \mathcal{J}}, \{U_j\}_{\forall j \in \mathcal{J}}]$ can also be determined by adopting a reinforcement learning-based approach. Specifically, in this paper, we explore the benefits of the log-linear-based reinforcement learning (RL) and we design two alternative algorithms, i.e., B-Logit, and Max-Logit. Both algorithms perform an exploration of the node's j strategy space by randomly selecting a strategy $\mathbf{s}_j$ with equal probability $P(\mathbf{s}_j) = \frac{1}{|\mathcal{J}|}$ and determine the corresponding payoff $U_j(\mathbf{s}_j, \mathbf{s}_{-j})$ that they experience. Then, at the learning phase, each target and collaborator update its strategy based on the probabilistic rules (7a)-(7b) for the B-Logit algorithm and (8a)-(8b), for the Max-Logit algorithm. The proposed log-linear RL algorithms are presented in Algorithm 3, where $\beta \in \mathbb{R}^+$ captures the learning parameter. It should be highlighted that the Max-Logit algorithm can determine the Pareto optimal Nash Equilibrium if it exists [30].

Algorithm 2 Asynchronous Best Response Dynamics (ABRD)

Input: $\mathbf{P}_u$, $\forall u \in \mathcal{U}$, $\hat{\mathbf{P}}_c$, $\forall c \in \mathcal{C}$
Output: $\boldsymbol{s}^*$
Initialization: $i = 0$, $Convergence = 0$, $\mathbf{s}^{i=0}$ randomly selected strategy.
while *Convergence* $== 0$ **do**
  $i = i + 1$;
  Select randomly a node $j \in \mathcal{J} = \mathcal{T} \cup \mathcal{C}$
  The selected node determines $\mathbf{s}_j^{*i}$ *(Eq. 5) and determines* $U_j(\mathbf{s}_j^{*i}, \mathbf{s}_{-j}^{*i})$ *(Eq. 4), given* $\mathbf{s_{-j}}^{i-1}$
  if $|U_j(\mathbf{s}_j^{*i}, \mathbf{s}_{-j}^{i-1}) - U_j(\mathbf{s}_j^{*i+1}, \mathbf{s}_{-j}^{i})| \leqslant \delta$, δ small positive number, $\forall j \in \mathcal{J}$ **then**
    $Convergence = 1$
  end if
end while

Algorithm 3 B-Logit (Max-Logit) Algorithm

Input: $\mathbf{P}_u$, $\forall u \in \mathcal{U}$, $\hat{\mathbf{P}}_c$, $\forall c \in \mathcal{C}$, β, I
Output: $\boldsymbol{s}^*$
Initialization: $i = 0$, $Convergence = 0$, $\mathbf{s}_j^{i=0}$, $\forall n \in \mathcal{J}$.
while *Convergence* $== 0$ **do**
  $i = i + 1$;
  Each target and collaborator node j *selects* $\mathbf{s}_j^{i'}$ *with equal probability* $\frac{1}{|\mathcal{S}_j|}$*, receives a payoff* $U_j(\mathbf{s}_j^{i'})$ *and updates* $\mathbf{s}_j^i$ *based on Eq. 7a*, 7b (Eq. 8a, 8b).
  The rest of the nodes keep their previous strategies, i.e., $\mathbf{s}_{-j}^i = \mathbf{s}_{-j}^{i-1}$.
  if $|\frac{\sum_{i=0}^{I} \sum_{\forall j \in \mathcal{J}} U_j^i}{I} - \sum_{\forall j \in \mathcal{J}} U_j^i| \leqslant \delta$, δ small positive number **then**
    $Convergence = 1$
  end if
end while

$$P(\mathbf{s}_j^i = \mathbf{s}_j^{i-1}) = \frac{e^{\beta U_j(\mathbf{s}_n^{i-1})}}{e^{\beta U_j(\mathbf{s}_j^{i-1})} + e^{\beta U_j(\mathbf{s}_j^{i'})}} \tag{7a}$$

$$P(\mathbf{s}_j^i = \mathbf{s}_j^{i'}) = \frac{e^{\beta U_j(\mathbf{s}_j^{i'})}}{e^{\beta U_j(\mathbf{s}_j^{i-1})} + e^{\beta U_j(\mathbf{s}_j^{i'})}} \tag{7b}$$

$$P(\mathbf{s}_j^i = \mathbf{s}_j^{i-1}) = \frac{e^{\beta U_j(\mathbf{s}_j^{i-1})}}{\max\{e^{\beta U_j(\mathbf{s}_j^{i-1})}, e^{\beta U_j(\mathbf{s}_j^{i'})}\}} \tag{8a}$$

$$P(\mathbf{s}_j^i = \mathbf{s}_j^{i'}) = \frac{e^{\beta U_j(\mathbf{s}_j^{i'})}}{\max\{e^{\beta U_j(\mathbf{s}_j^{i-1})}, e^{\beta U_j(\mathbf{s}_j^{i'})}\}} \tag{8b}$$

IV. Evaluation & Results

The performance evaluation of the proposed PNT solution is achieved via modeling and simulation. Specifically, in Section IV-A, the performance characteristics of the best response

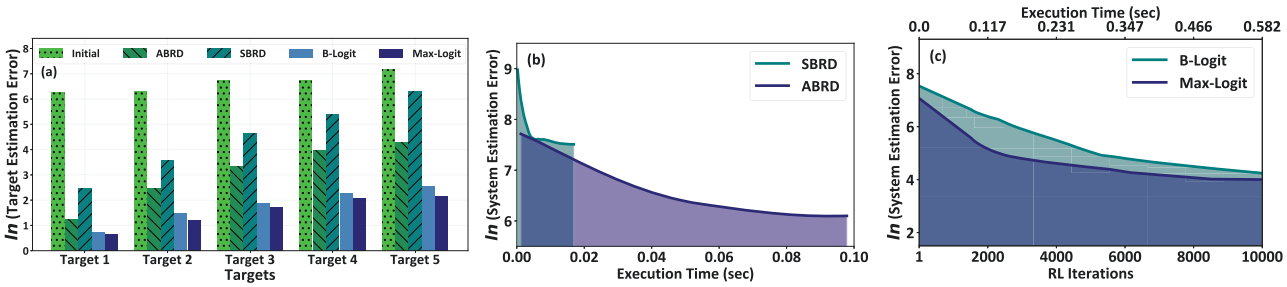

Fig. 3: Operational characteristics of the SBRD, ABRD, B-Logit, and Max-Logit algorithms.

Fig. 4: Trade-off among accuracy and complexity of the PNT on the air framework.

dynamics and reinforcement learning-based algorithms are demonstrated in terms of position and timing estimation error and convergence time. A detailed analysis of the trade-off between the accuracy of the proposed model and the corresponding complexity is provided in Section IV-B. The performance of the proposed model under the mobility use case scenario is analyzed in Section IV-C, while finally a comparative evaluation of the proposed approach against other state-of-the-art approaches is presented in Section IV-D. Unless otherwise explicitly stated, the simulation parameters that were used throughout our evaluation are listed as follows: $|\mathcal{U}| = 12$, $|\mathcal{C}| = 7$, $|\mathcal{T}| = 5$, $P = 2$ [W], $G_k^{trans} = 0$ [dB], $G_j^{rec} = 0$ [dB], $f_c = 400$ [MHz], $h_k^{trans} = 1.5$ [m], $h_j^{rec} = 1.5$ [m]. A Dell Tower Desktop with Intel i7 11700K 3.6GHz processor, 32 GB available RAM was used to conduct the evaluation.

A. Game-theoretic versus Reinforcement Learning-based Positioning, Navigation, and Timing on the Air

In this section, the operational characteristics of the BRD and RL algorithms are presented and compared in terms of the targets' estimation error, the system's estimation error, and the execution time in order to determine the Nash Equilibrium of the non-cooperative game. Specifically, Fig. 3a presents each target's estimation error for the four proposed algorithmic implementations, while they are also compared against the initial estimation error that the targets experience by implementing the Iterative Least Squares algorithm during the neighborhood identification process. Fig. 3b-3c present the overall system's estimation error of the targets and collaborators' position and timing for the SBRD and ABRD algorithms and the B-Logit and Max-Logit algorithms, respectively, as a function of the corresponding execution time of the algorithms.

The results reveal that the BRD algorithms are executed in a faster manner compared to the RL-based algorithms, given their deterministic decision-making process. On the other hand, the RL-based algorithms perform the exploration and learning phases in order to explore all the potential strategies that can minimize the estimation error experienced by the targets and collaborators, thus, the execution time of the RL-based algorithms is larger than the one of the BRD algorithms. Nevertheless, it is highlighted that the RL-based algorithms achieve lower system's estimation error (Fig. 3c) and lower estimation error for each target (Fig. 3a) given that they thoroughly explore all the available strategies. Moreover, it is observed that the Max-Logit algorithm achieves the best results in terms of both the targets (Fig. 3a) and the overall system's estimation error (Fig. 3c) given that it determines the Pareto optimal Nash Equilibrium of the non-cooperative game. Furthermore, the results reveal that the ABRD algorithm suffers from higher execution time due to the fact that the targets and collaborators update sequentially their best response strategies, compared to the SBRD algorithm. On the other hand, the SBRD algorithm, though it takes a shorter time to converge compared to the ABRD algorithm, it achieves a higher estimation error, as it suffers from the herding effect among the targets and collaborators, who update their strategies in a synchronous manner.

B. Accuracy and Complexity Analysis

In this section, the trade-off between the accuracy and the complexity of the four proposed algorithmic implementations is presented. Specifically, Fig. 4a presents the trade-off between the execution time and the system's estimation error as the convergence criterion δ of the SBRD and ABRD

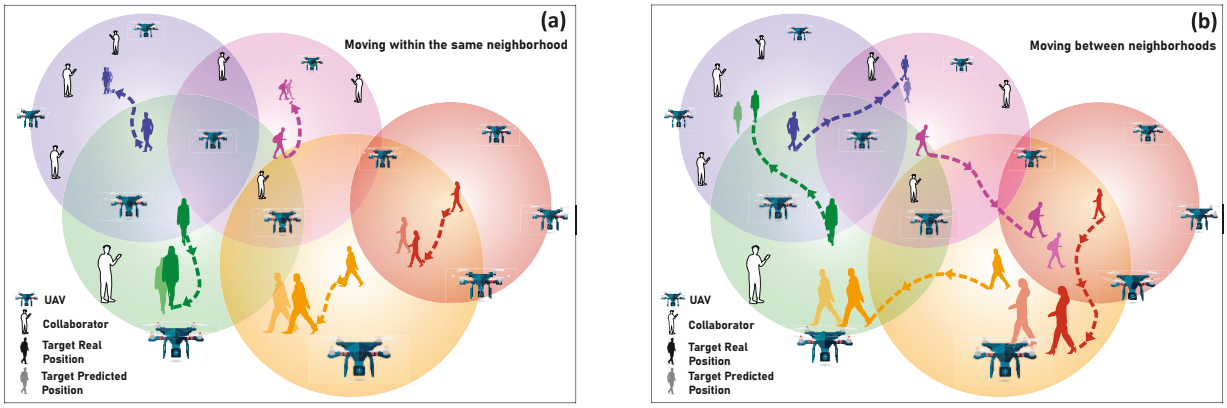

Fig. 5: Mobility analysis.

algorithms becomes stricter. Similarly, Fig. 4b presents the corresponding trade-off, as the learning parameter β increases, thus, allowing for a more thorough exploration phase.

The results reveal that as the convergence criterion δ becomes stricter for the BRD algorithms, their corresponding execution time increases, while the accuracy of the system improves in terms of experiencing lower estimation error for the targets' and collaborators' position and timing. By taking a closer look at the results, we also observe that the ABRD algorithm is significantly impacted by the stricter convergence criterion, in terms of its execution time, given that the targets and collaborators perform their best responses in a sequential manner (Fig. 4a). On the other hand, as the learning parameter of the RL-based algorithms increases, the targets and the collaborators explore more thoroughly their strategy space, thus, resulting in significantly lower systems estimation error (Fig. 4b). Also, the results reveal that under all the examined scenarios the Max-Logit algorithm achieves the lowest estimation error given that it determines the Pareto optimal Nash Equilibrium of the non-cooperative game.

C. Targets Mobility and Positioning, Navigation, and Timing

In this section, we present the benefits of the proposed PNT solution under a mobility use case scenario, where the target moves within the same neighborhood (Fig. 5a), or between neighborhoods (Fig. 5b). In the first case, i.e., Fig. 5a, the UAVs and the collaborators that support each target remain the same while the target moves. In the second case, where the target moves from one neighborhood to another one, the set of the UAVs and the collaborators that support its PNT services dynamically change over time. The results reveal that under both use case scenarios, the proposed PNT solution can accurately determine the position and timing of the target (shadowed presentation of the target) enabling its navigation in a real-time manner. However, it is observed that when the target moves between neighborhoods, the accuracy of the proposed PNT solution deteriorates, as the target changes dynamically the set of UAVs and collaborators that support its PNT services.

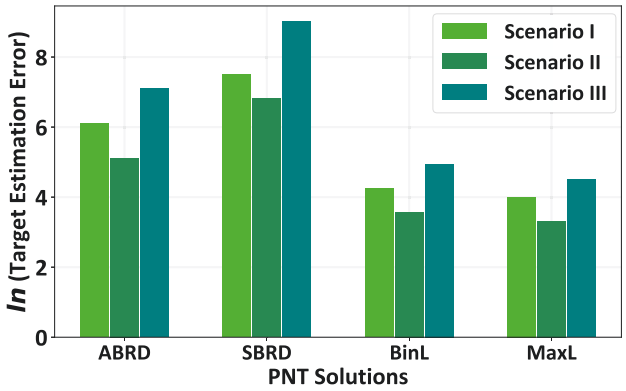

Fig. 6: Comparative analysis.

D. Comparative Evaluation

In this section, a comparative evaluation of: (i) Scenario I: the total target estimation error of the collaborative-based PNT solution (without the presence of collaborator nodes in the system), (ii) Scenario II: the total target estimation error of the collaborative-based PNT solution (with the presence of collaborator nodes in the system), (iii) Scenario III: the total target estimation error using the Iterative Least Squares algorithm which represents the vast majority of the current state-of-the-art approaches. Fig. 6 presents the estimation error of the overall system and of all the targets for the proposed PNT on the air framework versus the system's estimation error of the comparative scenario that performs the PNT without the existence of collaborators. The results show that the proposed PNT on the air framework achieves lower estimation error under all the four proposed algorithmic implementations. By taking a closer look at the results, we should highlight that in the proposed PNT on the air framework, the resulting error is lower compared to the scenario without collaborators, given that the collaborators significantly contribute in improving the overall system's error. Also, the results confirm that the Max-Logit algorithm achieves the lowest estimation error for the targets and the collaborators, given that it determines the Pareto optimal Nash Equilibrium of the non-cooperative game.

V. Conclusion and Future Work

In this paper, a novel collaborative-based PNT on the air framework is introduced by jointly exploiting the UAVs, collaborators, and targets towards accurately determining the position and timing of the two latter ones. An optimization problem of minimizing the position and timing estimation error of each target and collaborator, as well as of the overall system, has been formulated and addressed as a non-cooperative game among the targets and the collaborators. The existence of at least one NE point has been proven, while four alternative algorithmic implementation approaches have been introduced - following the principles of best response and reinforcement learning - in order to conclude to such an NE point. A detailed evaluation analysis is presented to demonstrate the benefits and tradeoffs of the proposed PNT solution.

Part of our current and future work includes the extension of the proposed model by considering a more holistic 3D network architecture consisting of targets, collaborators, UAVs, and high altitude platforms (HAPs), in order to exploit the full potential of the 3D networking environment in delivering accurate and reliable PNT services [32].

References

[1] X. Liu, J. Guan, R. Jiang, S. S. Ge, and B. Chen, "Finite-horizon urtss-based position estimation for urban vehicle localization," *IEEE Sensors Journal*, vol. 23, no. 4, pp. 4011–4021, 2023.

[2] L. Yang, N. Wu, B. Li, W. Yuan, and L. Hanzo, "Indoor localization based on factor graphs: A unified framework," *IEEE Internet of Things Journal*, vol. 10, no. 5, pp. 4353–4366, 2023.

[3] A. Pirnat, B. Bertalanič, G. Cerar, M. Mohorčič, M. Meža, and C. Fortuna, "Towards sustainable deep learning for wireless fingerprinting localization," in *ICC 2022 - IEEE International Conference on Communications*, 2022, pp. 3208–3213.

[4] J. Xue, J. Zhang, Z. Gao, and W. Xiao, "Enhanced wifi csi fingerprints for device-free localization with deep learning representations," *IEEE Sensors Journal*, vol. 23, no. 3, pp. 2750–2759, 2023.

[5] X. Zhu, T. Qiu, W. Qu, X. Zhou, M. Atiquzzaman, and D. O. Wu, "Bls-location: A wireless fingerprint localization algorithm based on broad learning," *IEEE Trans. on Mobile Computing*, vol. 22, no. 1, pp. 115–128, 2023.

[6] J. Paulavičius, S. Jardak, R. McConville, R. Piechocki, and R. Santos-Rodriguez, "Temporal self-supervised learning for rssi-based indoor localization," in *ICC 2022 - IEEE International Conference on Communications*, 2022, pp. 3046–3051.

[7] M. Heydariaan, H. Dabirian, and O. Gnawali, "Anguloc: Concurrent angle of arrival estimation for indoor localization with uwb radios," in *2020 16th DCOSS*, 2020, pp. 112–119.

[8] J. Wahlström, P. Porto Buarque de Gusmão, A. Markham, and N. Trigoni, "Map-aided navigation for emergency searches," in *2019 15th DCOSS*, 2019, pp. 25–32.

[9] K.-H. Lam, C.-C. Cheung, and W.-C. Lee, "Rssi-based lora localization systems for large-scale indoor and outdoor environments," *IEEE Trans. on Vehicular Technology*, vol. 68, no. 12, pp. 11 778–11 791, 2019.

[10] H. Guo, I. Quartey, and C. Green, "Rss-based localization using a single robot in complex environments," in *2022 18th DCOSS*, 2022, pp. 364–371.

[11] S. Rani, H. Babbar, P. Kaur, M. D. Alshehri, and S. H. Shah, "An optimized approach of dynamic target nodes in wireless sensor network using bio inspired algorithms for maritime rescue," *IEEE Trans. on Intelligent Transportation Systems*, vol. 24, no. 2, pp. 2548–2555, 2023.

[12] R. Su, Z. Gong, C. Li, and X. Shen, "Algorithm design and performance analysis of target localization using mobile underwater acoustic array networks," *IEEE Trans. on Vehicular Technology*, vol. 72, no. 2, pp. 2395–2406, 2023.

[13] I. Tallini, L. Iezzi, P. Gjanci, C. Petrioli, and S. Basagni, "Localizing autonomous underwater vehicles: Experimental evaluation of a long baseline method," in *2021 17th International Conference on Distributed Computing in Sensor Systems (DCOSS)*, 2021, pp. 443–450.

[14] Q. Zhang and W. Saad, "Semi-supervised learning for channel charting-aided iot localization in millimeter wave networks," in *2021 IEEE Global Communications Conference (GLOBECOM)*, 2021, pp. 1–6.

[15] R. Chen, X. Huang, Y. Zhou, Y. Hui, and N. Cheng, "Uhf-rfid-based real-time vehicle localization in gps-less environments," *IEEE Trans. on Intelligent Transportation Systems*, vol. 23, no. 7, pp. 9286–9293, 2022.

[16] X. An, X. Cui, S. Zhao, G. Liu, and M. Lu, "Efficient rigid body localization based on euclidean distance matrix completion for agv positioning under harsh environment," *IEEE Trans. on Vehicular Technology*, vol. 72, no. 2, pp. 2482–2496, 2023.

[17] E. Zhang and N. Masoud, "Increasing gps localization accuracy with reinforcement learning," *IEEE Trans. on Intelligent Transportation Systems*, vol. 22, no. 5, pp. 2615–2626, 2021.

[18] W. Zhang and W. Zhang, "An efficient uav localization technique based on particle swarm optimization," *IEEE Trans. on Vehicular Technology*, vol. 71, no. 9, pp. 9544–9557, 2022.

[19] A. Quan, C. Herrmann, and H. Soliman, "Project vulture: A prototype for using drones in search and rescue operations," in *2019 15th International Conference on Distributed Computing in Sensor Systems (DCOSS)*, 2019, pp. 619–624.

[20] C.-H. Lin, Y.-H. Fang, H.-Y. Chang, Y.-C. Lin, W.-H. Chung, S.-C. Lin, and T.-S. Lee, "Gcn-cnvps: Novel method for cooperative neighboring vehicle positioning system based on graph convolution network," *IEEE Access*, vol. 9, pp. 153 429–153 441, 2021.

[21] A. Chakraborty, K. M. Brink, and R. Sharma, "Cooperative relative localization using range measurements without a priori information," *IEEE Access*, vol. 8, pp. 205 669–205 684, 2020.

[22] Y. Zhao, X. Li, and M. Xia, "Cooperative localization in wireless powered communication network," in *ICC 2021 - IEEE International Conference on Communications*, 2021, pp. 1–6.

[23] J. S. Russell, M. Ye, B. D. O. Anderson, H. Hmam, and P. Sarunic, "Cooperative localization of a gps-denied uav using direction-of-arrival measurements," *IEEE Trans. on Aerospace and Electronic Systems*, vol. 56, no. 3, pp. 1966–1978, 2020.

[24] Y. Cao, S. Yang, Z. Feng, L. Wang, and L. Hanzo, "Distributed spatio-temporal information based cooperative 3d positioning in gnss-denied environments," *IEEE Trans. on Vehicular Technology*, vol. 72, no. 1, pp. 1285–1290, 2023.

[25] M. S. Siraj, A. B. Rahman, M. Diamanti, E. E. Tsiropoulou, S. Papavassiliou, and J. Plusquellic, "Orchestration of reconfigurable intelligent surfaces for positioning, navigation, and timing," in *MILCOM 2022 - 2022 IEEE Military Communications Conference (MILCOM)*, 2022, pp. 148–153.

[26] M. S. Siraj, A. B. Rahman, M. Diamanti, E. E. Tsiropoulou, and S. Papavassiliou, "Alternative positioning, navigation, and timing enabled by games in satisfaction form and reconfigurable intelligent surfaces," *IEEE Systems Journal*, pp. 1–12, 2023.

[27] 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 *Ad-hoc, Mobile, and Wireless Networks: 13th International Conference, ADHOC-NOW 2014, Benidorm, Spain, June 22-27, 2014 Proceedings 13*. Springer, 2014, pp. 114–127.

[28] V. Garg, "Radio propagation and propagation path-loss models," *Wireless Communications & Networking*, pp. 47–84, 2007.

[29] M. S. Hossain, N. Irtija, E. E. Tsiropoulou, J. Plusquellic, and S. Papavassiliou, "Reconfigurable intelligent surfaces enabling positioning, navigation, and timing services," in *ICC 2022 - IEEE International Conference on Communications*, 2022, pp. 4625–4630.

[30] D. Monderer and L. S. Shapley, "Potential games," *Games and economic behavior*, vol. 14, no. 1, pp. 124–143, 1996.

[31] G. Katsinis, E. E. Tsiropoulou, and S. Papavassiliou, "Joint resource block and power allocation for interference management in device to device underlay cellular networks: A game theoretic approach," *Mobile Networks and Applications*, vol. 22, pp. 539–551, 2017.

[32] M. S. Siraj, M. S. Hossain, R. Brown, E. E. Tsiropoulou, and S. Papavassiliou, "Incentives to learn: A location-based federated learning model," in *2022 Global Information Infrastructure and Networking Symposium (GIIS)*, 2022, pp. 40–45.