Article

Machine Learning Based Indoor Localization Under Shadowing Condition for P-NOMA VLC Systems

Affan Affan[1,‡], Hafiz M. Asif [2,‡,*] and Naser Tarhuni [2,‡]

[1] Department of Electrical and Computer Engineering, University of Louisville, KY, USA; affan.affan@louisville.edu
[2] Department of Electrical and Computer Engineering, Sultan Qaboos University, Oman; h.asif@squ.edu.om, tarhuni@squ.edu.om
* Correspondence: h.asif@squ.edu.om
‡ These authors contributed equally to this work.

Abstract: The communication link between the base station and the mobile sensor networks, such as multi-agent systems for collaborative tasks which include ground mobile robots and drones, is crucial for localization and success of tasks in indoor environments. The Power-domain Non-Orthogonal Multiple Access (P-NOMA) is an emerging multiplexing technique that enables the base station to accumulate signals for different agents using the same time-frequency channel. The environment information such as distance from the base station is required at the base station to calculate communication channel gains and allocate suitable signal power to each agent. The accurate estimate of the position for power allocation of P-NOMA in a dynamic environment is challenging due to changing location of the end-agent and shadowing. In this paper, we take advantage of the two-way Visible Light Communication (VLC) link to: (1) estimate the position of the end-agent in a real-time indoor environment based on the signal power received at the base station and (2) allocate resources using the Simplified Gain Ratio Power Allocation (S-GRPA) scheme with the look-up table method. In addition, we use the Euclidean Distance Matrix (EDM) to estimate the location of the end-agent whose signal was lost due to shadowing.

Keywords: Visible Light Communication; Machine Learning; SIC; NOMA; Localization; Shadowing

1. Introduction

Visible Light Communication (VLC) systems have been the focus of research and development for many years. Numerous VLC products have been developed and are commercially available such as Light Fidelity (LiFi) [1]. The basic components of VLC systems include Light Emitting Diodes (LEDs) as transmitters and Photo-Detectors (PDs) as the receivers. At the transmitter end, the information bits are converted into electrical signals to drive LEDs, while at the receiver end, the photons are received by PDs and converted back to information bits. VLC systems use the visible light spectrum instead of the Radio Frequency (RF) spectrum. The wide interference-free range of the visible light spectrum, $380 - 780\ nm$, is one of the main advantages of VLC systems over RF systems. VLC systems also provide a higher data rate [2,3]. VLC systems may not replace RF systems completely, however, these systems can be used as hybrid technologies to improve communication quality for different applications and environments [4].

VLC systems are used for both data communication and illumination. This leads to a wide range of applications [5]. Indoor environments are prime examples of the application of VLC systems due to the availability of pre-installed infrastructure for the transmitters and the receivers [6]. In [7], the authors proposed an indoor VLC system based on On-Off Keying (OOK) using four devices equipped with 3600 LEDs (60×60) which were installed in a room of dimensions $5.0\ m \times 5.0\ m \times 3.0\ m$. The authors concluded that the system can satisfy communication requirements for the indoor environment. In [8], a data rate of about

 doi:10.20944/preprints202305.0346.v1

200 *Mbps* was achieved using blue LEDs for an indoor environment. The laser-based white light emitting Surface Mount Device (SMD) platform combined with blue LEDs is proposed in [9], the proposed framework achieved a data rat of about 20 *Gbps* with $10 - 100$ times more brightness as compared to the conventional light bulbs.

Indoor positioning methods using VLC technologies have been reported in numerous studies. The features such as Received Signal Strength (RSS), Time of Arrival (ToA), and Angle of Arrival (AoA) are the most prominent when designing an indoor positioning system. The first, high-precision VLC-based indoor positioning system was proposed in [10]. The proposed algorithm uses RSS measurement to obtain user/object location with an accuracy of 0.4 m. In [11], AoA is used as a feature to find the user/object location with the accuracy of 0.1 m. The proposed framework used image sensors instead of PDs. The hybrid approach using a combination of RSS and AoA for improved communication and localization is proposed in [12,13]. In the aspect of cost-effectiveness, [14] proposed a two-stage framework. First, presented a dedicated analog sensor that is capable of being directly plugged into the microphone input of a computer or a mobile device such as a smartphone. The signal pattern, as well as the signal strength of a beacon, can both be decoded by it. Second, to decode the signal pattern, the use of rolling shutter cameras is proposed. This offers a viable answer to the problem of localizing hand-held devices that contain cameras.

Artificial Intelligence (AI) methods have been widely developed for the application of indoor positioning mainly due to high accuracy and easy deployment [15,16]. The machine learning-based classifiers are compared for localization in an indoor environment in [17]. It has been reported that the k-nearest neighbor (k-NN) algorithm performed the best among all other classifiers. It has been reported in [18], that support vector machine (SVM) outperformed logistic regression using the fingerprinting method for Bluetooth signals in an indoor environment. In [19] Principle Component Analysis (PCA) is used to improve the performance of SVM, K-NN, and random forest classifiers for indoor positioning. In [20], it is reported that the random forest classifier outperforms K-NN for WiFi-based indoor positioning. In [21], an enhance random forest algorithm is proposed for indoor positioning in real-world settings. Inspired by the literature, this paper compares the performance of SVM and random forest regression (RFR) for VLC-based indoor positioning.

VLC technologies have been used for a variety of autonomous systems such as vehicles, autonomous ground robots, and fixed robotic industrial grippers. In [22], an in-hospital transportation robot called HOSPI is developed and VLC technology is used to improve navigation and localization in addition to other navigation sensors for better autonomous control of the robot. The study explores the localization and autonomy using experimental and actual results in an actual hospital. In [23], the multi-frequency method with RSS is used to measure the distance of the robot from each LED that is installed above the robot in a plane that is parallel to the plane of the robot base in order to accomplish precise indoor positioning. This is accomplished by installing the LEDs in a plane that is perpendicular to the surface of the robot base. A multi-frequency technique is one in which each LED transmits its location ID at a frequency that is distinct from the others. In [24], VLC technology in addition to the Extended Kalman filter has been used for communication and localization in underwater robots for nuclear reactor inspection. In this research, a system is described that makes use of the modulated light signal both as a medium through which data can be sent and as a reference upon which to base the positioning of a mobile robot. Both of these functions are carried out by the same system. VLC technologies have been proven to be efficient and cost-effective. The underlying modulation and multiplexing techniques have constantly evolved to meet the needs of more complicated and dynamic environments.

Many different approaches, such as Orthogonal Frequency Division Multiple Access (OFDMA), Time Division Multiple Access (TDMA), and Code Division Multiple Access (CDMA), have been suggested in the research literature as methods for implementing multiple access in very low bit rate (VLC) systems [25–31]. MIMO-OFDM is utilized for

 doi:10.20944/preprints202305.0346.v1

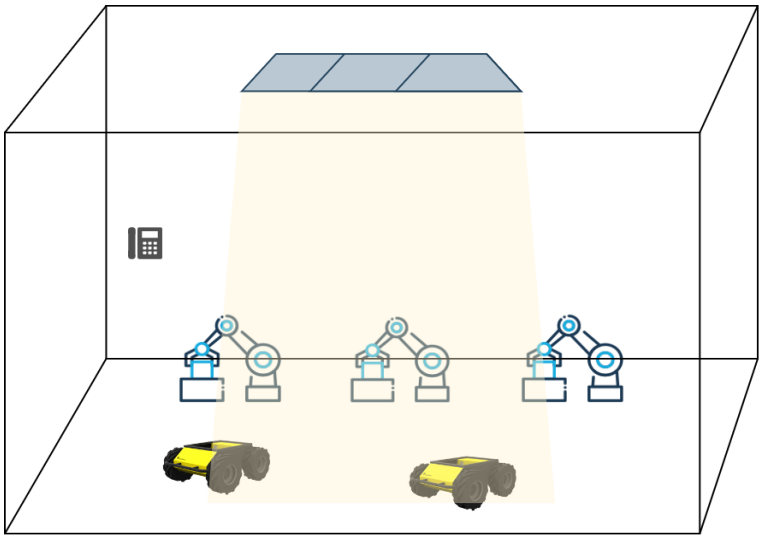

Figure 1. General indoor environment for collaborative tasks with LED arrays for transmission.

multiuser VLC systems in [32], and this is accomplished by giving each user their own unique carrier. In the same vein, CDMA is used in conjunction with OFDM to support multiple-user communication rather than sending a single carrier to each user. Similarly, Non-Orthogonal Multiple Access (NOMA) methods have been widely proposed in VLC systems to increase the number of users without compromising on performance [33]. In the power domain NOMA, the data of different agents are accumulated together using different power factors for each agent, and Successive Interference Cancellation (SIC) is used to retrieve the data at the receiver end. However, the resource allocation for the power domain NOMA is highly affected by the channel and environment.

The resource allocation in NOMA has been a focus of research for many years [34]. In [35], the subject of energy-efficient user planning and power optimization in NOMA wireless links is investigated to understand the trade-off that exists between the data rate effectiveness and the amount of energy that is consumed by NOMA. For the downlink NOMA heterogeneous network, energy-efficient user planning and power distribution techniques are presented for both perfect and imperfect CSI, respectively. In [36], the neural network-based resource allocation method is proposed for mobile users with mutual interference management. This research work also provides priorities and rate demands-based user scheduling methods to coordinate the access of heterogeneous users with limited radio resources. In [37], a low complexity power allocation scheme for NOMA-based indoor VLC systems which is called the Simplified Gain Ratio Power Allocation (S-GRPA) scheme is proposed. The Channel State Information (CSI) used for power allocation in NOMA is obtained through the look-up table method rather than calculation. In this research work, the location of the agent is received at the base station by a separate low-energy communication link. This approach suffers the loss of the signal-carrying position information. Inspired by this approach, we proposed to use machine learning algorithms to find the location of the agent for CSI for better dynamic resource allocation in NOMA for a collaborative indoor environment instead of the separate communication link. Additionally, to avoid loss of agent location due to shadowing or obstacles, we used Euclidian Distance Matrix (EDM) to obtain the location of agents using mutual distances of agents in the network.

The rest of the article is structured as follows. Section 2.2 provides a mathematical discussion on the S-GRPA for NOMA, followed by Section 2.1 which discusses the VLC channel. Section 2.3 & 2.4 discusses Random Forest Regression (RFR) and Support Vector Machine (SVM) algorithms. The section 2.5 discusses the Euclidian Distance Matrix (EDM) to obtain an agent's location in the event of signal loss due to shadowing and obstacles.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 May 2023 doi:10.20944/preprints202305.0346.v1

Section 3 presents results for indoor positioning and bit-error-rate (BER) for NOMA-based VLC system. The article is finally wrapped up with concluding remarks.

2. System Design

In this research work, a complete framework for resource allocation for power domain NOMA using S-GRPA, in the indoor environment for collaborative tasks, is proposed using visible light communication systems as shown in Fig. 1. For indoor positioning, we have compared the RFR and the SVM algorithms with Euclidian Distance Matrix (EDM). The complete system block diagram is shown in Fig. 2. The next section discusses the S-GRPA for NOMA with VLC channel model.

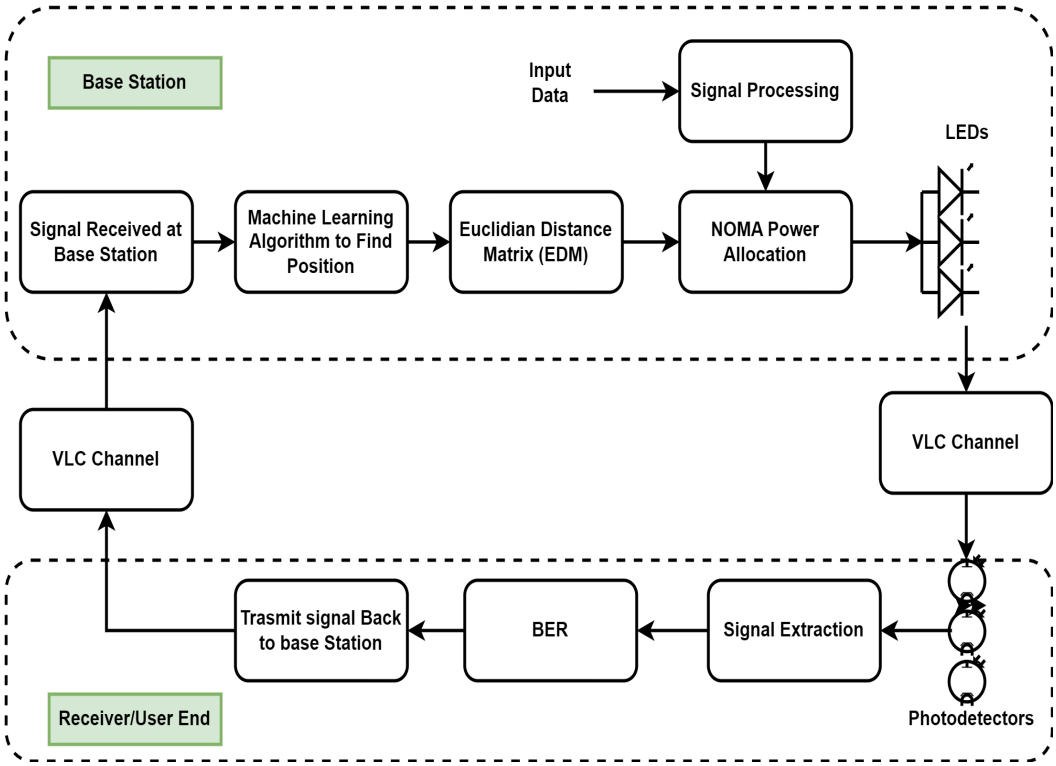

Figure 2. Block diagram for complete NOMA-based VLC system using S-GRPA and indoor positioning.

2.1. Visible Light Communication (VCL) Channel

LEDs, the transmitting devices in VLC, serve dual tasks by providing both light and data transmission. As can be seen in Fig. 3, the responsiveness of the VLC channel in an indoor setting is heavily dependent on the illumination intensity and the transmission power. The illumination intensity at a point in the Cartesian plane is given as follows [38]:

$$I(x,y) = \frac{I(0)cos^{m_l}(\theta)}{d^2 cos(\psi)}, \tag{1}$$

where $I(0)$ is the intensity of the central light source, m_l, is the order of Lambertian emission, d is the separation distance between the LEDs and the PDs, and θ and ψ are the irradiance and incidence angles, respectively. Lambertian emission, m_l, is described in the following order:

$$m_l = \frac{ln(2)}{ln(cos(\vartheta_{\frac{1}{2}}))}, \tag{2}$$

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 May 2023 doi:10.20944/preprints202305.0346.v1

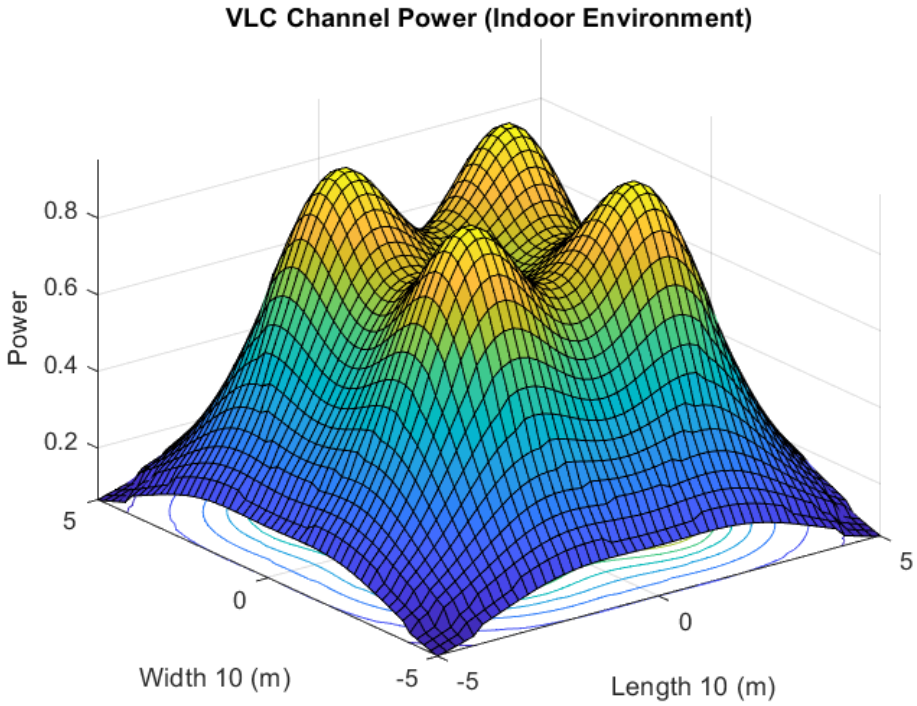

Figure 3. 3D view of power distribution of VLC channel in an indoor environment with four transmitting locations.

where $\vartheta_{\frac{1}{2}}$ is the angle at half illuminance of an LED. The signal power received at a particular PD is given as follows:

$$P_r = P_t.\left(\frac{m_l+1}{2\pi d^2}\right)cos^{m_l}(\theta).T_s(\psi).cos(\psi), \quad 0 \leq \psi \leq \psi_{con} \tag{3}$$

where $T_s(\psi)$ is the filter transmission, $g(\psi)$ is concentrator gain and ψ_{con} is the field of view of PD. P_t is transmitted signal power, it fades through Line of Sight (LoS) channel gain. The DC gain of the LoS path is given as follows:

$$h_d(0) = \begin{cases} \left(\frac{m_l+1}{2\pi d^2}\right)cos^{m_l}(\theta) & \psi \leq \psi_{con} \\ 0 & \psi > \psi_{con} \end{cases} \tag{4}$$

The distance, d, between the transmitter and receiver is the key factor in the allocation of signal power in NOMA-based communication systems. The next section discusses the mathematical basis for NOMA and resource allocation in this method.

2.2. Simplified Gain Ratio Power Allocation (S-GRPA) for NOMA

Several agents' data are combined using the power-domain NOMA, and data is then separated using SIC. NOMA is free of spectrum spreading or degraded SNR performance, in contrast to other multiple access methods such as OFDMA and CDMA [39]. Several agents using the same resource simultaneously in NOMA boosts system throughput, but each agent has a distinct power factor. In this study, we employ symbol-level NOMA. Fig. 5 illustrates the NOMA framework for three agents.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 May 2023 doi:10.20944/preprints202305.0346.v1

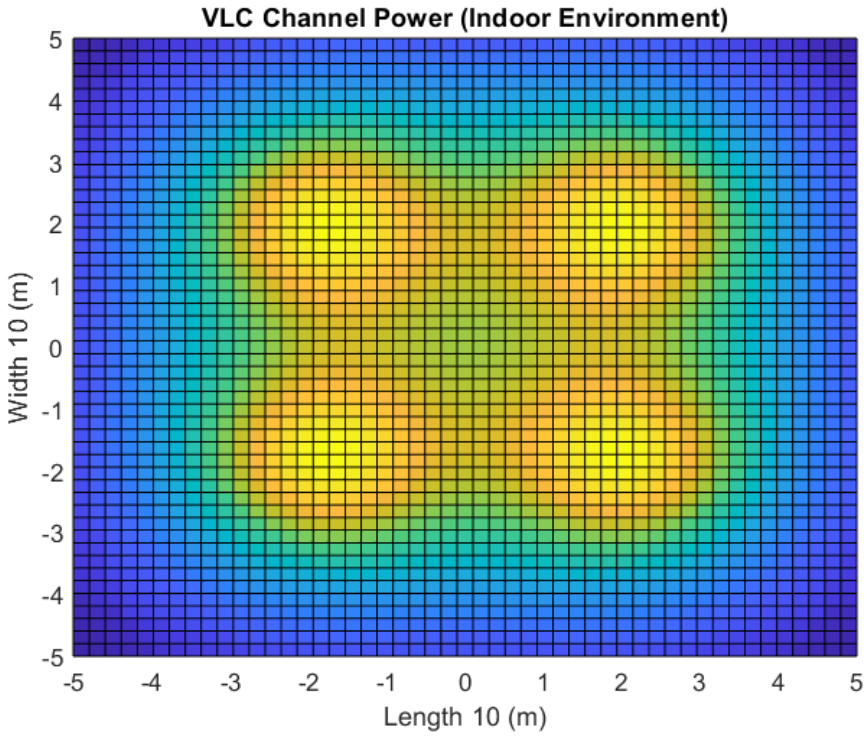

Figure 4. 2D view of power distribution of VLC channel in an indoor environment with four transmitting locations.

Because of the relative difference in distance from the base station, the agent near the cell edge gets a lot more power than the agent at the center of the cell. The received signal can be expressed as follows [39]:

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} H_1 \\ H_2 \end{bmatrix} \begin{bmatrix} p_1 x_1 & p_2 x_2 \end{bmatrix} + \begin{bmatrix} \eta_1 \\ \eta_2 \end{bmatrix} \tag{5}$$

where y_1 and y_2 are, the suppressed information from all agents, received at each agent. The information from agent-1 is separated and demodulated as follows at the receiver end:

$$\widehat{s}_1 = \frac{y_1}{p_1} \tag{6}$$

After agent-1's data has been retrieved, agent 2's data is recovered by SIC reducing agent-1's interference in the manner described below:

$$\widehat{z} = y_2 - p_1 \widehat{s}_1 \tag{7}$$

$$\widehat{s}_2 = \frac{\widehat{z}}{p_2} \tag{8}$$

This method can be expanded to a higher number of agents. As the power factor, p, for each agent depends on its location. The power distribution of the VLC system is shown in Fig.4, it can be seen that overall power distribution can be divided into zones, $n = 1, 2, \cdots, N$, with radius, r_N. Here r_n refers to the radius of each zone. Furthermore, zone N has the least illumination intensity, and zone 1 has the highest illumination intensity. The n^{th} region can be defined as follows [37]:

$$\begin{cases} [0, r_n] & n = 1 \\ [r_{n-1}, r_n] & 1 < n < N \\ [r_{n-1}, r_e] & n = N \end{cases} \tag{9}$$

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 May 2023 doi:10.20944/preprints202305.0346.v1

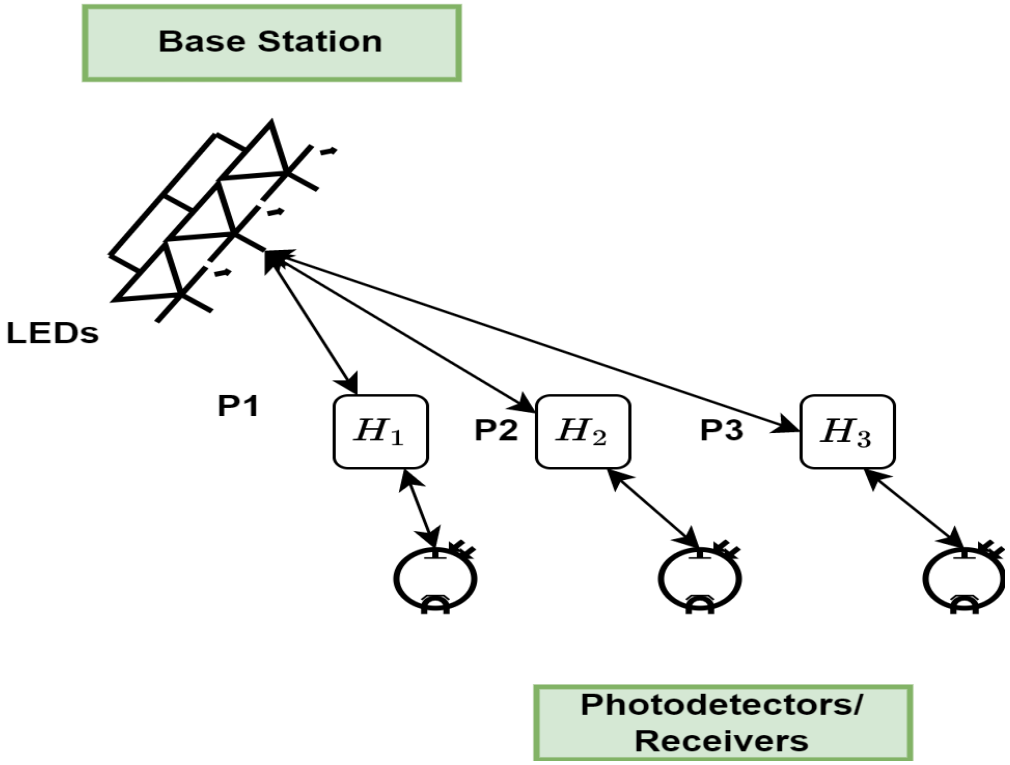

Figure 5. Illustration of NOMA method for VLC systems

for

$$r_n = \sqrt{\frac{nr_e^2}{N}}, n = 1, 2, \cdots, N. \tag{10}$$

where r_e is the maximum radius of the last zone. The relationship between the k^{th} agent and the $k-1^{th}$ agent in terms of power allocation is described as follows [37]:

$$p_k = \left(\frac{H_1}{H_N}\right)^k \times p_{k-1}, \tag{11}$$

Here, the information on channel gain, H, is received by the indoor positioning algorithm and stored in the look-up table.

2.3. Random Forest Regression Algorithm

The idea of aggregating random decision trees was initially presented in the research papers[40–42]. This idea is at the heart of the RFR technique. The following is the formulation of the problem statement for the RFR algorithm:

Problem 1. *Calculate the value of the non-parametric regression function denoted by* $f(x) = E[y|\mathbf{x}] \rightarrow y = f(x)$*, where* p *is the dimension of the input vector and* $\mathbf{x} \in \mathbb{R}^p$ *is the input vector used to estimate the output,* $y \in \mathbb{R}$*.*

Using the training data set, the RFR method predicts the function $f_e(x)$, which is similar to the actual regression function $f(x)$, and then compares the two. $T_n = \{[|y_1|\mathbf{x_1}], [|y_2|\mathbf{x_2}], \cdots, [|y_n|\mathbf{x_n}]\}$ is a mathematical expression. The error function can be written as: $E[f_e(x) - f(x)]^2 \rightarrow 0$, as $n \rightarrow \infty$. The random forest methodology employs N

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 May 2023 doi:10.20944/preprints202305.0346.v1

different types of regression trees. The input vector, **x**, is evaluated on each tree in such a way that the projected value for the i^{th} tree is given as follows [41]:

$$f_e(\mathbf{x}, \theta_i, T_n) = \sum_{j \in T_n(\theta_i)} \frac{\mathbf{x_j} \in A_m(\mathbf{x}, \theta_i, T_n) y}{M_m(\mathbf{x}, \theta_i, T_n)} \tag{12}$$

here, $\theta_1, \theta_2, \cdots, \theta_N$ are the independent random variables associated with each regression tree. $T_n(\theta_i)$ are the data points selected prior to the construction of trees. $A_m(\mathbf{x}, \theta_i, T_n)$ is the zone containing **x** and $M_m(\mathbf{x}, \theta_i, T_n)$ is the number of data points fall into $A_m(\mathbf{x}, \theta_i, T_n)$.

2.4. Support Vector Machine Algorithm

As with the RFR algorithm, the purpose of the Support Vector Machine (SVM) regression algorithm is to map the input **x** to the desired output y. Instead of minimizing the discrepancy between the estimated and true regression functions, f_e and f, respectively, SVM penalizes the final result so that it can be used in linear regression. $y = w.x + b$. Let y be the true output, and inside an area defined as $y \pm \epsilon$, SVM will make its predictions. The discrepancy between actual output, y, and expected output, z, is denoted by the $|z_i - y_i| < \epsilon$. Two slack variables are added as a penalty if the projected output is outside the region $y \pm \epsilon$. ζ^+ indicates that the projected output lies above the region of $y + \epsilon$, whereas ζ^- indicates that the predicted output lies below the region of $y - \epsilon$. The following is the error function for the SVM regression algorithm: [43–47]:

$$min \quad C \sum_{i=1}^{L} (\zeta_i^+ + \zeta_i^-) + \frac{1}{2} \|w\|^2, \tag{13}$$

subject to:

$$\begin{aligned} &\zeta^+ \geq 0, \\ &\zeta^- \geq 0, \\ &y_i \leq z_i + \epsilon + \zeta^+ \\ &y_i \geq z_i + \epsilon + \zeta^- \end{aligned} \tag{14}$$

C is the tunable variable that controls the penalty on slack variables and ϵ. To solve (13), following Lagrange multipliers are introduced [43,47]:

$$\begin{aligned} &\alpha_i^+ \geq 0, \alpha_i^- \geq 0, \\ &\mu_i^+ \geq 0, \mu_i^- \geq 0. \end{aligned} \tag{15}$$

This transforms (13) in following:

$$\begin{aligned} L_p = &C \sum_{i=1}^{L} (\zeta_i^+ + \zeta_i^-) + \frac{1}{2} \|w\|^2 - \sum_{i=1}^{L} (\mu_i^+ \zeta_i^+ + \mu_i^- \zeta_i^-) \\ &- \sum_{i=1}^{L} \alpha_i^+ (\epsilon + \zeta_i^+ + z_i - y_i) - \sum_{i=1}^{L} \alpha_i^- (\epsilon + \zeta_i^- - z_i + y_i) \end{aligned} \tag{16}$$

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 May 2023 doi:10.20944/preprints202305.0346.v1

Now differentiating (16) to 0, with respect to each variable as below [43,47]:

$$\begin{aligned}
\frac{\partial L_p}{\partial w} &= 0 \Rightarrow w = \sum_{i=1}^{L}(\alpha_i^+ - \alpha_i^-)\mathbf{x_i} \\
\frac{\partial L_p}{\partial b} &= 0 \Rightarrow \sum_{i=1}^{L}(\alpha_i^+ - \alpha_i^-) = 0 \\
\frac{\partial L_p}{\partial \zeta_i^+} &= 0 \Rightarrow C = \alpha_i^+ + \mu_i^+ \\
\frac{\partial L_p}{\partial \zeta_i^-} &= 0 \Rightarrow C = \alpha_i^- + \mu_i^-
\end{aligned} \tag{17}$$

The Dual form of the Primary L_p can be defined as follows:

$$\min_{\alpha^+,\alpha^-} \begin{bmatrix} \sum_{i=1}^{L}(\alpha_i^+ - \alpha_i^-)y_i - \epsilon\sum_{i=1}^{L}(\alpha_i^+ - \alpha_i^-) \\ -\frac{1}{2}\sum_{i,j}(\alpha_i^+ - \alpha_i^-)(\alpha_i^+ - \alpha_i^-)\mathbf{x_i}\cdot\mathbf{x_j} \end{bmatrix} \tag{18}$$

subject to:

$$\begin{aligned}
&0 \le \alpha_i^+ \le C, \\
&0 \le \alpha_i^- \le C, \\
&\sum_{i=1}^{L}(\alpha_i^+ - \alpha_i^-) = 0.
\end{aligned} \tag{19}$$

The predicted output can be written as [47]:

$$z = \sum_{i=1}^{L}(\alpha_i^+ - \alpha_i^-)\mathbf{x}_i.\mathbf{x} + b. \tag{20}$$

This concludes the regression algorithms to obtain the location of the agent based on power received at the receiver. However, in real-world settings, the direct receiver might not be able to receive the agent's signal due to obstacles or the signal is highly disrupted due to noise. Therefore, for localization in collaborative indoor environments, the distance geometry problem plays an important role by obtaining the location of the particular agent using the mutual distances of all agents in the network. The next section discusses the distance geometry problem using Euclidian Distance Matrix (EDM).

2.5. The Distance Geometry Problem

The VLC systems for indoor environments with dynamic agents often face the loss of signal due to shadowing because of obstacles. To assist the localization in the event of signal loss in a multiagent collaborative environment, the Distance Geometry Problem (DGP) is used in conjunction with the machine learning algorithms discussed in the previous subsection. The objective of the DGP is to find the location of agents using their mutual distances. It is important to mention that agents do not only communicate with the base station during collaborative tasks but also communicate with each other, this provides extra information about the environment that agents transmit to the base station.

The mutual distances of agents are provided in the form of EDM. The solution to DGP for N agents in dimension, d, is a matrix, $S \in \mathbb{R}^{d\times N} = [s_1, s_2, \ldots, s_N]$, here s_i are the coordinates for i^{th} agent. The mutual distance between agents can be referred as $D \in \mathbb{R}^{N\times N} = [d_{ij}]$, here d_{ij} is the distance from j^{th} agent to i^{th} agent. To generate an initial estimate of the matrix $\hat{S}$, multi-dimensional scaling is used. The estimated point cloud, $\hat{S}$ can be mapped to the actual matrix, S, using rigid transformation in absolute coordinates. For this transformation, the Procrustes Analysis is performed which is spectral factorization.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 May 2023 doi:10.20944/preprints202305.0346.v1

Table 1. Simulation Parameters

Parameters	Value
Dimension	6.3 x 2.5 x 2 m^3
Tx-Rx (LED arrays-PDs on agents) Configuration	4X4
No. of LED Arrays	4
No. of LEDs in single Array	60
Total Power	20 watts
Semi-angle at half power	70 degrees
PD field of view	60 degrees
Refractive Index	1.5
No. of Agents	5

For this process, it is assumed that the location of some agents, $N_a < N$ is known, these agents are referred to as anchors. To find the location of missing agents in the network, DGP can be stated as static.

The static DGP consists of three stages to obtain the matrix, S from the matrix, D. The first stage is to obtain a Grammian matrix, $G \in \mathbb{R}^{N\times N} = S^T S$, it has one-one relation with matrix, D. The Grammian matrix, G, can be obtained by following the optimization problem [48,49]:

$$\begin{aligned} &\underset{G}{\text{minimize}} && ||\tilde{D} - W \circ \mathcal{K}(G)||^2 \\ &\text{subject to} && \\ & && G \succeq 0; \quad G1 = 0; \quad \text{Rank}(G) \leq d, \end{aligned}$$

where $\mathcal{K}(.)$ is a function, which maps the Grammian matrix, G, to the matrix, D. W is referred to as a binary mask matrix, the entries with zero values in this matrix show the missing measurement of the received power and $\circ$ is the Hadamard product. The next stage is to obtain the matrix, S, using the Grammian matrix, G. The estimate, $\hat{S}$ can be obtained by the Singular Value Decomposition (SVD) method given that matrices, G, and S hold the mathematical relation, $G = S^T S$.

As mentioned earlier that matrix, $\hat{S}$ can be mapped to the actual matrix, S, using rigid transformation along with the information of N_a anchors. By denoting the columns of the matrix, S, related to anchors as S_a and assuming Y_a refers to the same columns in $\hat{S}$. Furthermore, to make the S_a and Y_a centered at the origin, assume that S_a and Y_a are the translated version of S and Y. The transformation $\mathcal{R}$ can be defined as follows:

$$\mathcal{R} = \underset{\mathcal{Q}:\mathcal{Q}\mathcal{Q}^T - I}{\arg\min} ||\mathcal{Q}\bar{S}_a - \bar{Y}_a||_F^2$$

The actual matrix, S can be calculated as follows:

$$S = \mathcal{R}(\hat{S} - s_{a,c}1^T) + y_{a,c}1^T, \tag{21}$$

where $s_{a,c}$ and $y_{a,c}$ refer to the centroids of S_a and Y_a respectively.

This concludes the mathematical basis for the proposed framework for resource allocation using precise positioning of the agent in a collaborative indoor environment using VLC systems. The next section discusses the performance of indoor positioning using SVM and RFR algorithms, and bit-error-rate (BER) for NOMA-based VLC system.

3. Results

In the context of this research, a rectangular indoor environment that is analogous to Fig.1 is being studied. The rectangular environment has a width of 6.3 m, a depth of

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 May 2023 doi:10.20944/preprints202305.0346.v1

2.5 m, and a height of 3 m. These dimensions represent the area of the environment. Each Cartesian point has a zone that corresponds to it that is represented by a square cell that is 0.3 m^2 in size. The detailed simulation parameters are shown in Table 1. The purpose of

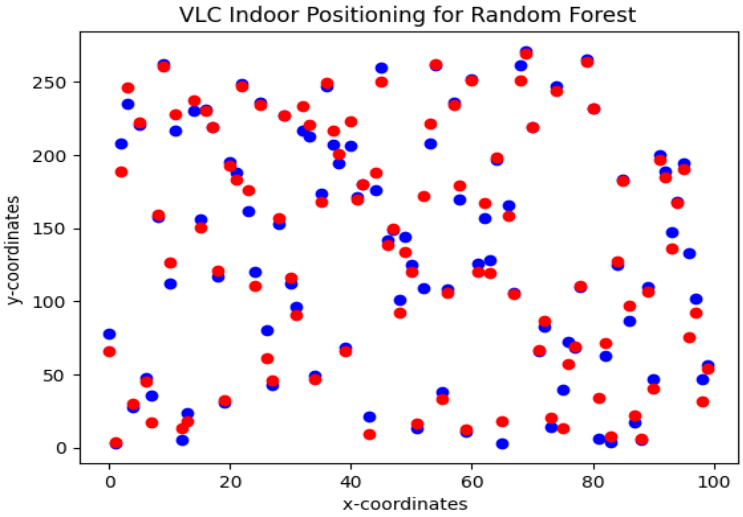

Figure 6. VLC indoor positioning using Random forest regression algorithm. actual data points (blue), predicted data points (red)

this exercise is to make an estimation of the Cartesian location within the square cell region and use this information to calculate channel gain for resource allocation for NOMA.

To obtain the location of agents, the SVM and RFR algorithms are trained using RSS and AoA features at the receiver end. Due to 4×4 Tx-Rx (LED arrays-PDs on agents) configuration for each agent, each feature has a dimension of $\mathbb{R}^{4\times4}$. The dataset is collected in MATLAB using the VLC channel. Each agent has a region of confidence of 0.25 m for its location. Fig. 6 shows the estimated and actual Cartesian coordinates of agents. For this simulation, the agent is considered to be dynamic, therefore, different locations are estimated in these simulations for more robustness in the algorithm. The red dots are actual coordinates and the blue dots are estimated values. For RFR, 100 decision trees are used in this simulation. The RFR shows an accuracy of 93.6% with an estimation error of 0.19 ± 0.22.

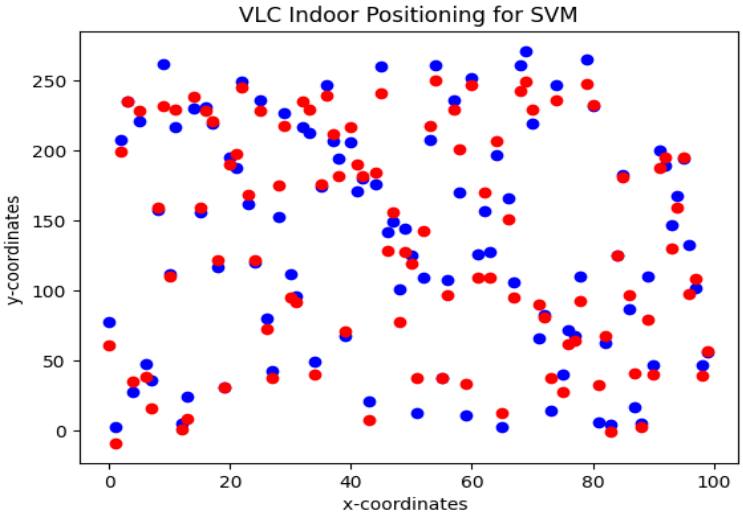

Figure 7. VLC indoor positioning using SVM regression algorithm. actual data points (blue), predicted data points (red)

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 May 2023 doi:10.20944/preprints202305.0346.v1

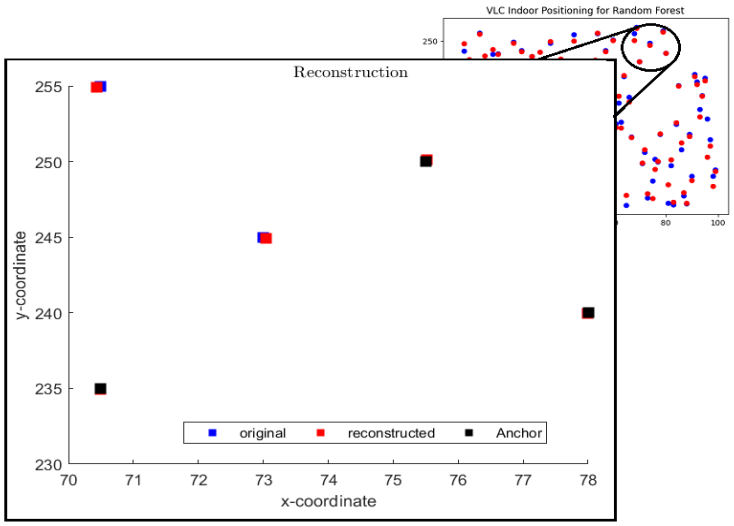

Figure 8. VLC indoor positioning using EDM with RFR, when agents signal is not received at the base due to obstacle. Actual data points (blue), predicted data points using EDM (red), and predicted data points using RFR (black) called anchors.

Fig. 7 shows the results of the estimated location using the SVM algorithm. For SVM algorithm radial basis function is used to solve the error function. The SVM shows an accuracy of 84% with an estimation error of 0.3 ± 0.2. According to the findings of the statistical analysis, the mean error for the SVM regression method is higher than the mean error for the random forest regression technique. Therefore, RFR is used as a positioning algorithm in conjunction with S-GRPA for power allocation in a NOMA-based multi-agent VLC system.

Fig. 8 shows a scenario of five agents where the signal of two agents is not received at the base station due to obstacle/shadowing. In this scenario, the location of the other three agents is predicted by the RFR algorithm and used as anchors for EDM as shown by black data points. The red square boxes in Fig. 8 show the estimated position of two remaining agents using EDM. it can be seen that EDM with RFR predicted the locations precisely. Here the predicted position of anchors (black boxes) is considered to be accurate during the process of EDM.

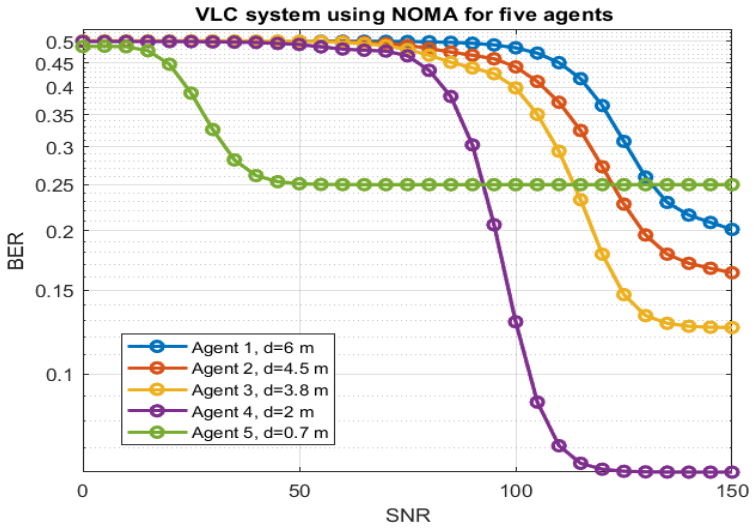

Figure 9. BER results of five agents for NOMA-based LoS VLC system.

 doi:10.20944/preprints202305.0346.v1

By combining the RFR algorithm with S-GRPA for power factor allocation in the VLC system, Fig. 9 shows the BER curves for five agents at different locations. The results consist of a Signal-to-Noise Ratio (SNR) between 0 to 150. The blue curve shows agent-1 at a distance of 6 m inches which is the farthest agent in the NOMA setup. The orange curve shows the BER rate for agent-2 at a distance of 4.5 m in an indoor environment. The yellow curve shows the BER for agent-3 at a distance of 3.8 m. The purple and green curves show the BER curves for agent-4 and agent-5 at distances of 2 and 0.7 m respectively. As expected the BER for the closest agent shows sharp and quick decay however, the farthest agent shows slow and high BER values. The power factor for each agent is allocated based on the location found by the RFR algorithm. It can be seen that S-GRPA with RFR is able to allocate appropriate power factors to each agent based on the location in an indoor environment.

4. Conclusion

Indoor localization and task accomplishment depend on the communication link between the base station and mobile sensor networks, such as multi-agent systems for collaborative tasks that involve ground mobile robots and drones. In this research work, a power allocation method using the precise location of the agent is proposed for the power-domain non-orthogonal multiple access (P-NOMA). P-NOMA enables the base station to suppress signals for different agents using the same time-frequency channel, it requires the information of the channel for better power allocation to each agent. The channel is highly affected by the distance between the end-agent and the base station. In this research work machine learning algorithm is proposed to find the location of the agent using Received Signal Strength (RSS) and Angle of Arrival (AoA). The location provided by the machine learning algorithm is used to determine the channel gain and the appropriate power factor is allocated to each agent based on simplified gain ratio power allocation (S-GRPA). It is shown by simulations that Random Forest Regression (RFR) performed better to obtain location as compared to Support Vector Machine (SVM). In addition, the Euclidian Distance matrix is used to find the location of the agent, if the signal is not received from a particular agent, based on the mutual distances of agents in the network. The complete framework is tested for five agents using the VLC channel. The S-GRPA with RFR algorithm was able to assign the appropriate power factors to each user based on its location. This works provides bases for dynamic power allocation for multiuser VLC systems, The future work aims to test this method in real-world experiments.

Author Contributions: Conceptualization, Hafiz Asif and Affan; methodology, all the authors; software, Affan; validation, all the authors; formal analysis, all the authors; investigation, Hafiz Asif and Naser Tarhuni; resources, Hafiz Asif; writing—original draft preparation, Affan; writing—review and editing, all the authors; supervision, Hafiz Asif and Naser Tarhuni; project administration, Hafiz Asif; funding acquisition, Hafiz Asif and Naser Tarhuni.

Funding: This research work is funded by Internal Grant (IG), Sultan Qaboos University under the project code IG/ENG/ECED/21/03.

Acknowledgments: The authors would like to acknowledge the SQU support in successfully executing this research work. The authors would like to thank Dr. Talha Manzoor of the Center of Water Informatics (WIT), Lahore University of Management Sciences (LUMS), Lahore, Pakistan for the valuable discussions on EDM. The authors would also like to thank Dr. Muhammad Zaigham Abbas Shah of Dept. of Electronic Engineering, Mehran University of Engineering and Technology, Jamshoro, Sindh, Pakistan for the valuable discussion on machine learning algorithms.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Swami, K.T.; Moghe, A.A. A review of LiFi technology. In Proceedings of the 2020 5th IEEE international conference on recent advances and innovations in engineering (ICRAIE). IEEE, 2020, pp. 1–5.
2. Basnayaka, D.A.; Haas, H. Hybrid RF and VLC systems: Improving user data rate performance of VLC systems. In Proceedings of the 2015 IEEE 81st vehicular technology conference (VTC Spring). IEEE, 2015, pp. 1–5.

3. Ramadhani, E.; Mahardika, G. The technology of lifi: A brief introduction. In Proceedings of the IOP conference series: materials science and engineering. IOP Publishing, 2018, Vol. 325, p. 012013.
4. Alfattani, S. Review of LiFi technology and its future applications. *Journal of Optical Communications* **2021**, *42*, 121–132.
5. Matheus, L.E.M.; Vieira, A.B.; Vieira, L.F.; Vieira, M.A.; Gnawali, O. Visible light communication: concepts, applications and challenges. *IEEE Communications Surveys & Tutorials* **2019**, *21*, 3204–3237.
6. O'Brien, D.; Le Minh, H.; Zeng, L.; Faulkner, G.; Lee, K.; Jung, D.; Oh, Y.; Won, E.T. Indoor visible light communications: challenges and prospects. *Free-Space Laser Communications VIII* **2008**, *7091*, 60–68.
7. Komine, T.; Nakagawa, M. Fundamental analysis for visible-light communication system using LED lights. *IEEE transactions on Consumer Electronics* **2004**, *50*, 100–107.
8. Grubor, J.; Langer, K.; Walewski, J.; Randel, S. High-speed wireless indoor communication via visible light. *ITG fachbericht* **2007**, *198*, 203.
9. Lee, C.; Islim, M.S.; Videv, S.; Sparks, A.; Shah, B.; Rudy, P.; McLaurin, M.; Haas, H.; Raring, J. Advanced LiFi technology: laser light. In Proceedings of the Light-emitting devices, materials, and applications XXIV. SPIE, 2020, Vol. 11302, pp. 116–123.
10. Li, L.; Hu, P.; Peng, C.; Shen, G.; Zhao, F. Epsilon: A visible light based positioning system. In Proceedings of the 11th {USENIX} Symposium on Networked Systems Design and Implementation ({NSDI} 14), 2014, pp. 331–343.
11. Kuo, Y.S.; Pannuto, P.; Hsiao, K.J.; Dutta, P. Luxapose: Indoor positioning with mobile phones and visible light. In Proceedings of the Proceedings of the 20th annual international conference on Mobile computing and networking, 2014, pp. 447–458.
12. Prince, G.B.; Little, T.D. A two phase hybrid RSS/AoA algorithm for indoor device localization using visible light. In Proceedings of the 2012 IEEE Global Communications Conference (GLOBECOM). IEEE, 2012, pp. 3347–3352.
13. Yang, S.H.; Kim, H.S.; Son, Y.H.; Han, S.K. Three-dimensional visible light indoor localization using AOA and RSS with multiple optical receivers. *Journal of Lightwave Technology* **2014**, *32*, 2480–2485.
14. Liu, M.; Qiu, K.; Che, F.; Li, S.; Hussain, B.; Wu, L.; Yue, C.P. Towards indoor localization using visible light communication for consumer electronic devices. In Proceedings of the 2014 IEEE/RSJ International Conference on Intelligent Robots and Systems. IEEE, 2014, pp. 143–148.
15. Nessa, A.; Adhikari, B.; Hussain, F.; Fernando, X.N. A survey of machine learning for indoor positioning. *IEEE access* **2020**, *8*, 214945–214965.
16. Roy, P.; Chowdhury, C. A survey of machine learning techniques for indoor localization and navigation systems. *Journal of Intelligent & Robotic Systems* **2021**, *101*, 63.
17. Bozkurt, S.; Elibol, G.; Gunal, S.; Yayan, U. A comparative study on machine learning algorithms for indoor positioning. In Proceedings of the 2015 International Symposium on Innovations in Intelligent SysTems and Applications (INISTA). IEEE, 2015, pp. 1–8.
18. Sthapit, P.; Gang, H.S.; Pyun, J.Y. Bluetooth based indoor positioning using machine learning algorithms. In Proceedings of the 2018 IEEE International Conference on Consumer Electronics-Asia (ICCE-Asia). IEEE, 2018, pp. 206–212.
19. Salamah, A.H.; Tamazin, M.; Sharkas, M.A.; Khedr, M. An enhanced WiFi indoor localization system based on machine learning. In Proceedings of the 2016 International conference on indoor positioning and indoor navigation (IPIN). IEEE, 2016, pp. 1–8.
20. Jedari, E.; Wu, Z.; Rashidzadeh, R.; Saif, M. Wi-Fi based indoor location positioning employing random forest classifier. In Proceedings of the 2015 international conference on indoor positioning and indoor navigation (IPIN). IEEE, 2015, pp. 1–5.
21. Maung, N.A.M.; Lwi, B.Y.; Thida, S. An enhanced RSS fingerprinting-based wireless indoor positioning using random forest classifier. In Proceedings of the 2020 International Conference on Advanced Information Technologies (ICAIT). IEEE, 2020, pp. 59–63.
22. Murai, R.; Sakai, T.; Kawano, H.; Matsukawa, Y.; Kitano, Y.; Honda, Y.; Campbell, K.C. A novel visible light communication system for enhanced control of autonomous delivery robots in a hospital. In Proceedings of the 2012 IEEE/SICE International Symposium on System Integration (SII). IEEE, 2012, pp. 510–516.
23. Sharifi, H.; Kumar, A.; Alam, F.; Arif, K.M. Indoor localization of mobile robot with visible light communication. In Proceedings of the 2016 12th IEEE/ASME International Conference on Mechatronic and Embedded Systems and Applications (MESA). IEEE, 2016, pp. 1–6.
24. Rust, I.C.; Asada, H.H. A dual-use visible light approach to integrated communication and localization of underwater robots with application to non-destructive nuclear reactor inspection. In Proceedings of the 2012 IEEE International Conference on Robotics and Automation. IEEE, 2012, pp. 2445–2450.
25. Affan, A.; Khan, U.; Asif, H.M.; Raahemifar, K. Multiuser visible light communication system using hybrid OFDM-PWM. In Proceedings of the 2020 International Symposium on Networks, Computers and Communications (ISNCC). IEEE, 2020, pp. 1–6.
26. Chen, D.; Wang, J.; Jin, J.; Lu, H.; Feng, L. A CDMA system implementation with dimming control for visible light communication. *Optics Communications* **2018**, *412*, 172–177.
27. Matsushima, T.K.; Sasaki, S.; Kakuyama, M.; Yamasaki, S.; Murata, Y.; Teramachi, Y. A visible-light communication system using optical CDMA with inverted MPSC. In Proceedings of the The Sixth International Workshop on Signal Design and Its Applications in Communications. IEEE, 2013, pp. 52–55.
28. Qiu, Y.; Chen, S.; Chen, H.H.; Meng, W. Visible light communications based on CDMA technology. *IEEE Wireless Communications* **2017**, *25*, 178–185.

29. Affan, A.; Mumtaz, S.; Asif, H.M.; Musavian, L. Performance analysis of orbital angular momentum (oam): A 6g waveform design. *IEEE Communications Letters* **2021**, *25*, 3985–3989.
30. Asif, H.M.; Affan, A.; Tarhuni, N.; Raahemifar, K. Deep Learning-Based Next-Generation Waveform for Multiuser VLC Systems. *Sensors* **2022**, *22*, 2771.
31. Lian, J.; Brandt-Pearce, M. Multiuser visible light communication systems using OFDMA. *Journal of Lightwave Technology* **2020**, *38*, 6015–6023.
32. Wang, Q.; Wang, Z.; Dai, L. Multiuser MIMO-OFDM for visible light communications. *IEEE Photonics Journal* **2015**, *7*, 1–11.
33. Ren, H.; Wang, Z.; Du, S.; He, Y.; Chen, J.; Han, S.; Yu, C.; Xu, C.; Yu, J. Performance improvement of NOMA visible light communication system by adjusting superposition constellation: a convex optimization approach. *Optics express* **2018**, *26*, 29796–29806.
34. Islam, S.M.R.; Zeng, M.; Dobre, O.A.; Kwak, K.S. Resource Allocation for Downlink NOMA Systems: Key Techniques and Open Issues. *IEEE Wireless Communications* **2018**, *25*, 40–47. https://doi.org/10.1109/MWC.2018.1700099.
35. Zhang, H.; Fang, F.; Cheng, J.; Long, K.; Wang, W.; Leung, V.C. Energy-efficient resource allocation in NOMA heterogeneous networks. *IEEE Wireless Communications* **2018**, *25*, 48–53.
36. Liu, M.; Song, T.; Gui, G. Deep cognitive perspective: Resource allocation for NOMA-based heterogeneous IoT with imperfect SIC. *IEEE Internet of Things Journal* **2018**, *6*, 2885–2894.
37. Zhao, Q.; Jiang, J.; Wang, Y.; Du, J. A low complexity power allocation scheme for NOMA-based indoor VLC systems. *Optics Communications* **2020**, *463*, 125383.
38. Ghassemlooy, Z.; Popoola, W.; Rajbhandari, S. *Optical wireless communications: system and channel modelling with Matlab®*; CRC press, 2019.
39. Yan, C.; Harada, A.; Benjebbour, A.; Lan, Y.; Li, A.; Jiang, H. Receiver design for downlink non-orthogonal multiple access (NOMA). In Proceedings of the 2015 IEEE 81st vehicular technology conference (VTC Spring). IEEE, 2015, pp. 1–6.
40. Breiman, L. Random forests. *Machine learning* **2001**, *45*, 5–32.
41. Biau, G.; Scornet, E. A random forest guided tour. *Test* **2016**, *25*, 197–227.
42. Scornet, E. On the asymptotics of random forests. *Journal of Multivariate Analysis* **2016**, *146*, 72–83.
43. Stitson, M.; Weston, J.; Gammerman, A.; Vovk, V.; Vapnik, V. Theory of support vector machines. *University of London* **1996**, *117*, 188–191.
44. Chen, P.H.; Fan, R.E.; Lin, C.J. A study on SMO-type decomposition methods for support vector machines. *IEEE transactions on neural networks* **2006**, *17*, 893–908.
45. Fan, R.E.; Chen, P.H.; Lin, C.J.; Joachims, T. Working set selection using second order information for training support vector machines. *Journal of machine learning research* **2005**, *6*.
46. Cristianini, N.; Shawe-Taylor, J.; et al. *An introduction to support vector machines and other kernel-based learning methods*; Cambridge university press, 2000.
47. Fletcher, T. Support vector machines explained. *Tutorial paper* **2009**, pp. 1–19.
48. Tabaghi, P.; Dokmanić, I.; Vetterli, M. Kinetic Euclidean distance matrices. *IEEE Transactions on Signal Processing* **2019**, *68*, 452–465.
49. Dokmanic, I.; Parhizkar, R.; Ranieri, J.; Vetterli, M. Euclidean distance matrices: essential theory, algorithms, and applications. *IEEE Signal Processing Magazine* **2015**, *32*, 12–30.