

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, Louisville, KY 40292, USA; affan.affan@louisville.edu
- ² Department of Electrical and Computer Engineering, Sultan Qaboos University, Al-Khoud, Muscat 123, Oman; tarhuni@squ.edu.om
- * Correspondence: h.asif@squ.edu.om

Abstract: The localization of agents for collaborative tasks is crucial to maintain the quality of the communication link for successful data transmission between the base station and agents. 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 the 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 using machine learning algorithms 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. The simulation results show that the machine learning algorithm is able to provide an accuracy of 0.19 m and allocate power to the agent.

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) and Laser Diodes (LDs) as transmitters and Photo-Detectors (PDs) as the receivers [2,3]. 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 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 [4,5]. 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 [6].

VLC systems are used for both data communication and illumination. This leads to a wide range of applications [7]. Indoor environments are prime examples of the application of VLC systems due to the availability of pre-installed infrastructure for the transmitters [8]. In [9], 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 of [9] concluded that the system can satisfy communication requirements, such as the high data rate of 10 Gbps can be achieved



Citation: Affan, A.; Asif, H.M.; Tarhuni, N. Machine-Learning-Based Indoor Localization under Shadowing Condition for P-NOMA VLC Systems. *Sensors* 2023, 23, 5319. https://doi.org/10.3390/s23115319

Academic Editor: Chi-Wai Chow

Received: 3 May 2023 Revised: 28 May 2023 Accepted: 1 June 2023 Published: 3 June 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/). using LEDs with OOK and intersymbol interference is dependent on the data rate and the FOV of the receiver in the indoor environment. In [10], a data rate of about 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 [11], and the proposed framework achieved a data rate 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 [12]. The proposed algorithm uses RSS measurement to obtain the user/object location with an accuracy of 0.4 m. In [13], AoA is used as a feature to find the user/object location with an 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 [14,15]. In the aspect of cost-effectiveness, Ref. [16] proposed a two-stage framework. First, they 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 [17,18]. The machine learning-based classifiers are compared for localization in an indoor environment in [19]. It has been reported that the k-nearest neighbor (k-NN) algorithm performed the best among all other classifiers. It has been reported in [20] that support vector machine (SVM) outperformed logistic regression using the fingerprinting method for Bluetooth signals in an indoor environment. In [21], Principle Component Analysis (PCA) is used to improve the performance of SVM, K-NN, and random forest classifiers for indoor positioning. In [22], it is reported that the random forest classifier outperforms K-NN for WiFi-based indoor positioning. In [23], an enhance random forest algorithm is proposed for indoor positioning in real-world settings. The proposed machine learning algorithms have proven to be efficient; however, a comprehensive analysis for a multi-agent system that also considers the signal loss due to obstacles is missing a link. Therefore, firstly this paper compares the performance of SVM and Random Forest Regression (RFR) for VLC-based indoor positioning and used Euclidian Distance Matrix (EDM) to find the location of the agent in the scenario of signal loss due to obstacles. Secondly, we complete the framework with adaptive resource allocation for multi-agent systems based on obtained positions.

VLC technologies have been used for a variety of autonomous systems such as vehicles, autonomous ground robots, and fixed robotic industrial grippers. In [24], 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 [25], 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 [26], VLC technology in addition to the Extended Kalman filter has been used for communication and localization in underwater robots for nuclear reactor inspection. VLC-based systems have proven to be efficient in the area of positioning for multiple applications; however, these proposed studies often require high-cost, sophisticated setups for high accuracy. Therefore, 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/agent. Both of these functions are carried out within the same framework.

Moreover, the underlying modulation and multiplexing techniques for VLC systems 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 high bit rate VLC systems [27–34]. MIMO-OFDM is utilized for multiuser VLC systems in [35], 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 [36]. 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 [37]. In [38], 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 Channel State Information (CSI), respectively. In [39], the neural network-based resource allocation method is proposed for mobile users with mutual interference management. This research work also provides priorities and rate demand-based user scheduling methods to coordinate the access of heterogeneous users with limited radio resources. In [40], 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 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 lowenergy 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 the loss of agent locations due to shadowing or obstacles, we used EDM to obtain the location of agents using mutual distances of agents in the network.

In this research work, we use machine learning algorithms such as RFR and SVM with minimum features of RSS and AoA for an indoor positioning framework using VLC technologies. The positioning algorithm is combined with the S-GRPA scheme for the adaptive resource allocation in a P-NOMA-based multi-agent communication network. Furthermore, we use the Euclidean Distance Matrix (EDM) to estimate the position of the agent whose signal was lost due to shadowing. The simulation results show that RFR-based positioning algorithm is able to provide an accuracy of 0.19 m and allocate power to the agent.

The rest of the article is structured as follows. Section 2.1 discusses the VLC channel followed by a mathematical discussion on the S-GRPA for NOMA in Section 2.2. Sections 2.3 and 2.4 discusses Random Forest Regression (RFR) and Support Vector Machine (SVM) algorithms. Section 2.5 discusses the Euclidean Distance Matrix (EDM) to obtain an agent's location in the event of signal loss due to shadowing and obstacles. Section 3 presents results for indoor positioning and the bit-error-rate (BER) for NOMAbased VLC systems. 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 Figure 1. For indoor positioning, we have compared the RFR and the SVM algorithms with the Euclidean Distance Matrix (EDM). The complete system block diagram is shown in Figure 1. The next section discusses the S-GRPA for NOMA with the VLC channel model.



Figure 1. Block diagram for complete NOMA-based VLC system using S-GRPA and indoor positioning.

2.1. Visible Light Communication (VLC) Channel

LEDs, the transmitting devices in VLC, serve dual tasks by providing both light and data transmission. As can be seen in Figure 2, 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 [41]:

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

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

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:

$$P_r = P_t \cdot \left(\frac{m_l + 1}{2\pi d^2}\right) \cos^{m_l}(\theta) \cdot T_s(\psi) \cdot \cos(\psi), \quad 0 \le \psi \le \psi_{con}$$
(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, and it fades through Line of Sight (LoS) channel gain. The DC gain of the LoS path is given as follows:

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



Figure 2. Three-dimensional view of relative power distribution of VLC channel in an indoor environment with four transmitting locations.

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 the data are then separated using SIC. NOMA is free of spectrum spreading and has low degradation SNR performance, in contrast to other multiple access methods such as OFDMA and CDMA, due to the high peak-to-average power ratio, low complex receiver design, and moderate latency [42,43]. 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. Figure 3 illustrates the NOMA framework for three agents. Because of the relative difference in distance from the base station, the agent near the cell edge receives a lot more power than the agent at the center of the cell. The received signal can be expressed as follows [42]:

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} H_1 \\ H_2 \\ H_3 \end{bmatrix} \begin{bmatrix} p_1 x_1 + p_2 x_2 + p_3 x_3 \end{bmatrix} + \begin{bmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \end{bmatrix}.$$
 (5)

The term $p_1x_1 + p_2x_2 + p_3x_3$ shows the infusion of data of each user at the base station and is transmitted to each user via the communication channel. These infused data experience different channel effects, such as H_1 , H_2 , H_3 , as shown by Figure 3. The terms y_1 , y_2 , and y_3

show the data received at each agent after passing through the channel. 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 have been retrieved, agent 2's data are recovered by SIC reducing agent-1's interference in the manner described below:

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

$$\widehat{s}_2 = \frac{\widehat{z}}{p_2}.$$
(8)



Figure 3. Illustration of NOMA method for VLC systems.

This method can be expanded to a higher number of agents. The power factor, p, for each agent depends on its location. The information about the position of agents in a multi-agent system can help to assign appropriate power factors to the agents for the P-NOMA-based VLC system. The power distribution of the VLC system is shown in Figure 2, where it can be seen that overall power distribution can be divided into zones, $n = 1, 2, \dots, N$, with radius, r_N . Here, r_n refers to the radius of each zone. Furthermore, zone N has the lowest illumination intensity and zone 1 has the highest. The nth region can be defined as follows [40]:

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

$$r_n = \sqrt{\frac{nr_\ell^2}{N}}, n = 1, 2, \cdots, N.$$
(10)

for

where r_e is the maximum radius of the last zone. The relationship between the *k*th agent and the k - 1th agent in terms of power allocation is described as follows [40]:

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

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 [44–46]. 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. $T_n = \{[|y_1|\mathbf{x_1}], [|y_2|\mathbf{x_2}], \dots, [|y_n|\mathbf{x_n}]\}$ 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* different types of regression trees. The input vector, \mathbf{x} , is evaluated on each tree in such a way that the projected value for the *i*th tree is given as follows [45]:

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

Here, $\theta_1, \theta_2, \dots, \theta_N$ 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 \mathbf{x} and $M_m(\mathbf{x}, \theta_i, T_n)$ is the number of data points that 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 [47–51]:

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

subject to:

$$\begin{split} \zeta^{+} &\geq 0, \\ \zeta^{-} &\geq 0, \\ y_{i} &\leq z_{i} + \epsilon + \zeta^{+}, \\ y_{i} &\geq z_{i} + \epsilon + \zeta^{-}. \end{split} \tag{14}$$

C is the tunable variable that controls the penalty on slack variables and ϵ . To solve (13), the following Lagrange multipliers are introduced [47,51]:

$$\begin{array}{l}
\alpha_{i}^{+} \geq 0, \, \alpha_{i}^{-} \geq 0, \\
\mu_{i}^{+} \geq 0, \, \mu_{i}^{-} \geq 0.
\end{array}$$
(15)

This transforms (13) in the following:

$$L_{p} = C \sum_{i=1}^{L} (\zeta_{i}^{+} + \zeta_{i}^{-}) + \frac{1}{2} ||w||^{2} - \sum_{i=1}^{L} (\mu_{i}^{+} \zeta_{i}^{+} + \mu_{i}^{-} \zeta_{i}^{-}) - \sum_{i=1}^{L} \alpha_{i}^{+} (\epsilon + \zeta_{i}^{+} + z_{i} - y_{i}) - \sum_{i=1}^{L} \alpha_{i}^{-} (\epsilon + \zeta_{i}^{-} - z_{i} + y_{i}).$$
(16)

Now differentiate (16) to 0, with respect to each variable as below [47,51]:

$$\frac{\partial L_p}{\partial w} = 0 \Rightarrow w = \sum_{i=1}^{L} (\alpha_i^+ - \alpha_i^-) \mathbf{x}_i,$$

$$\frac{\partial L_p}{\partial b} = 0 \Rightarrow \sum_{i=1}^{L} (\alpha_i^+ - \alpha_i^-) = 0,$$

$$\frac{\partial L_p}{\partial \zeta_i^+} = 0 \Rightarrow C = \alpha_i^+ + \mu_i^+,$$

$$\frac{\partial L_p}{\partial \zeta_i^-} = 0 \Rightarrow C = \alpha_i^- + \mu_i^-.$$
(17)

The dual form of the primary L_p can be defined as follows:

$$\min_{\alpha^{+},\alpha^{-}} \left[\sum_{i=1}^{L} (\alpha_{i}^{+} - \alpha_{i}^{-}) y_{i} - \epsilon \sum_{i=1}^{L} (\alpha_{i}^{+} - \alpha_{i}^{-}) \\ -\frac{1}{2} \sum_{i,j} (\alpha_{i}^{+} - \alpha_{i}^{-}) (\alpha_{i}^{+} - \alpha_{i}^{-}) \mathbf{x}_{i} \cdot \mathbf{x}_{j} \right],$$
(18)

subject to:

$$0 \le \alpha_{i}^{+} \le C, \\ 0 \le \alpha_{i}^{-} \le C, \\ \sum_{i=1}^{L} (\alpha_{i}^{+} - \alpha_{i}^{-}) = 0.$$
(19)

The predicted output can be written as [51]:

$$z = \sum_{i=1}^{L} (\alpha_i^+ - \alpha_i^-) \mathbf{x}_i \cdot \mathbf{x} + b.$$
(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 because 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, and 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, $S \in \mathbb{R}^{d \times N} = [s_1, s_2, ..., s_N]$, 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. For this process, it is assumed that the location of some agents, $N_a < N$ is known, and 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 matrix *S* from matrix *D*. The first stage is to obtain a Grammian matrix, $G \in \mathbb{R}^{N \times N} = S^T S$, as it has a one–one relation with matrix *D*. Grammian matrix *G* can be obtained by following the optimization problem [52,53]:

minimize _G	$ \tilde{D} - W \circ$	$\mathcal{K}(G) ^2$	
subject to			
	$G \succeq 0; G$	G1 = 0;	$\operatorname{Rank}(G) \leq d$

where $\mathcal{K}(.)$ is a function, which maps Grammian matrix *G* to 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 matrix *S* using 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, 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 matrix Srelated to anchors as S_a and assuming Y_a refers to the same columns in \hat{S} . Furthermore, to make S_a and Y_a centered at the origin, assume that S_a and Y_a are the translated version of Sand Y. The transformation \mathcal{R} can be defined as follows:

$$\mathcal{R} = rgmin_{\mathcal{Q}:\mathcal{Q}\mathcal{Q}^T-I} ||\mathcal{Q}ar{S}_a - ar{Y}_a||_F^2$$

The actual matrix *S* can be calculated as follows:

$$S = \mathcal{R}(\hat{S} - s_{a,c}1^{T}) + y_{a,c}1^{T},$$
(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 of a width of 6.3 m, a depth of 2.5 m, and a height of 3 m is being studied. 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 this exercise is to make an estimation of the Cartesian location within the square cell region and use this information to calculate the channel gain for resource allocation for NOMA. The location of agents consists of 2D Cartesian coordinates; however, the environment is three-dimensional. The dimensions $6.3 \times 2.5 \times 2 \text{ m}^3$ refer to length \times width \times height. The base station is considered to be mounted at 2 m height from the ground and agents are moving in 2D plane on the floor.

Table 1. Simulation parameters.

Parameters	Value
Dimension	$6.3\times2.5\times2m^3$
Tx-Rx (LED arrays-PDs on agents) Configuration	4 imes 4
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

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. Figure 4 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 \pm 0.22 m.



Figure 4. VLC indoor positioning using Random forest regression algorithm. Actual data points (blue), and predicted data points (red).

Figure 5 shows the results of the estimated location using the SVM algorithm. For the SVM algorithm, the 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.



Figure 5. VLC indoor positioning using SVM regression algorithm. Actual data points (blue), and predicted data points (red).

Table 2 shows the comparison of the proposed RFR and SVM algorithm for VLC-based indoor positioning with other state-of-the-art algorithms. In this table, the Epsilon, Plugo, and Luxapose are algorithms for indoor positioning using VLC technology. The WiDeep is the Wifi-based deep learning algorithm for indoor positioning. It can be seen that RFR with minimum features provides better accuracy than Epsilon, Plugo and WiDeep, and SVM, However, Luxapose is still better than the other algorithms due to more sophisticated hardware and a higher number of features. Therefore, it can be deduced that with minimum features, RFR provides better accuracy and easy deployment.

Figure 6 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 Figure 6 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.

Algorithm/Study	Accuracy
Proposed RFR method	0.19 m
Proposed SVM method	0.3 m
Epsilon [12]	4 m
Plugo [54]	0.33 m
WiDeep [55]	1.21 m
Luxapose [13]	0.1 m

Table 2. I enormance comparison	Table 2.	Performance	comparison
--	----------	-------------	------------



Figure 6. VLC indoor positioning using EDM with RFR, when agent 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.

By combining the RFR algorithm with S-GRPA for power factor allocation in the VLC system, Figure 7 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 high power factor is allocated to this user. 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. Agent-5 has been allocated low power based on the distance. As expected, the BER for the high-power agent shows sharp and quick decay; however, the low-power agent shows slow and high BER values. The higher BER shows the signal suffers higher noise during signal propagation from the base station to the agent. At a given time instance, the power allocation is performed based on the location of the agent. However, the value of the power factor based on distance from the base station is already stored in a look table using the simplified gain ratio power allocation (S-GRPA) method. 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.



Figure 7. BER results of five agents for NOMA-based LoS VLC system.

4. Conclusions

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 powerdomain non-orthogonal multiple access (P-NOMA). P-NOMA enables the base station to suppress signals for different agents using the same time-frequency channel, and 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, a 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 the 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 the RFR algorithm was able to assign the appropriate power factors to each user based on its location. This work 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, H.M.A. and A.A.; methodology, all the authors; software, A.A.; validation, all the authors; formal analysis, all the authors; investigation, H.M.A. and N.T.; resources, H.M.A.; writing—original draft preparation, A.A.; writing—review and editing, all the authors; supervision, H.M.A. and N.T.; project administration, H.M.A.; funding acquisition, H.M.A. and N.T. All authors have read and agreed to the published version of the manuscript.

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

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to privacy and ethical reasons.

Acknowledgments: The authors would like to acknowledge the SQU support in successfully executing this research work. The authors would like to thank 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 Muhammad Zaigham Abbas Shah of Department 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), Jaipur, India, 1–3 December 2020; pp. 1–5.
- Janjua, B.; Oubei, H.M.; Retamal, J.R.D.; Ng, T.K.; Tsai, C.T.; Wang, H.Y.; Chi, Y.C.; Kuo, H.C.; Lin, G.R.; He, J.H.; et al. Going beyond 4 Gbps data rate by employing RGB laser diodes for visible light communication. *Opt. Express* 2015, 23, 18746–18753. [CrossRef] [PubMed]
- Kumar, N.; Lourenco, N.R. Led-based visible light communication system: A brief survey and investigation. J. Eng. Appl. Sci 2010, 5, 296–307. [CrossRef]
- 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), Glasgow, UK, 11–14 May 2015; pp. 1–5.
- Ramadhani, E.; Mahardika, G. The technology of lifi: A brief introduction. In Proceedings of the IOP Conference Series: Materials Science and Engineering; IOP Publishing: Bristol, UK, 2018; Volume 325, p. 012013.

- 6. Alfattani, S. Review of LiFi technology and its future applications. J. Opt. Commun. 2021, 42, 121–132. [CrossRef]
- Matheus, L.E.M.; Vieira, A.B.; Vieira, L.F.; Vieira, M.A.; Gnawali, O. Visible light communication: Concepts, applications and challenges. *IEEE Commun. Surv. Tutor.* 2019, 21, 3204–3237. [CrossRef]
- 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 Commun. VIII* 2008, 7091, 60–68.
- Komine, T.; Nakagawa, M. Fundamental analysis for visible-light communication system using LED lights. *IEEE Trans. Consum. Electron.* 2004, 50, 100–107. [CrossRef]
- Grubor, J.; Langer, K.; Walewski, J.; Randel, S. High-speed wireless indoor communication via visible light. *ITG Fachber.* 2007, 198, 203.
- 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: Washington, DC, USA, 2020; Volume 11302, pp. 116–123.
- 12. 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), Seattle, WA, USA, 2–4 April 2014; pp. 331–343.
- Kuo, Y.S.; Pannuto, P.; Hsiao, K.J.; Dutta, P. Luxapose: Indoor positioning with mobile phones and visible light. In Proceedings of the 20th Annual International Conference on Mobile Computing and Networking, New York, NY, USA, 7–11 September 2014; pp. 447–458.
- 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), Anaheim, CA, USA, 3–7 December 2012; pp. 3347–3352.
- 15. 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. *J. Light. Technol.* 2014, *32*, 2480–2485. [CrossRef]
- 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, Chicago, IL, USA, 14–18 September 2014; pp. 143–148.
- 17. Nessa, A.; Adhikari, B.; Hussain, F.; Fernando, X.N. A survey of machine learning for indoor positioning. *IEEE Access* 2020, *8*, 214945–214965. [CrossRef]
- 18. Roy, P.; Chowdhury, C. A survey of machine learning techniques for indoor localization and navigation systems. *J. Intell. Robot. Syst.* **2021**, *101*, 63. [CrossRef]
- 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), Madrid, Spain, 2–4 September 2015; pp. 1–8.
- 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), Jeju, Republic of Korea, 24–26 June 2018; pp. 206–212.
- 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), Alcala de Henares, Spain, 4–7 October 2016; pp. 1–8.
- 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), Banff, AB, Canada, 13–16 October 2015; pp. 1–5.
- 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), Yangon, Myanmar, 4–5 November 2020; pp. 59–63.
- 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), Fukuoka, Japan, 16–18 December 2012; pp. 510–516.
- 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), Auckland, New Zealand, 29–31 August 2016; pp. 1–6.
- 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, St. Paul, MN, USA, 14–18 May 2012; pp. 2445–2450.
- 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), Montreal, QC, Canada, 20–22 October 2020; pp. 1–6.
- Chen, D.; Wang, J.; Jin, J.; Lu, H.; Feng, L. A CDMA system implementation with dimming control for visible light communication. Opt. Commun. 2018, 412, 172–177. [CrossRef]

- 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 Sixth International Workshop on Signal Design and Its Applications in Communications, Tokyo, Japan, 27 October–1 November 2013; pp. 52–55.
- Qiu, Y.; Chen, S.; Chen, H.H.; Meng, W. Visible light communications based on CDMA technology. *IEEE Wirel. Commun.* 2017, 25, 178–185. [CrossRef]
- 31. Affan, A.; Mumtaz, S.; Asif, H.M.; Musavian, L. Performance analysis of orbital angular momentum (oam): A 6g waveform design. *IEEE Commun. Lett.* 2021, 25, 3985–3989. [CrossRef]
- Asif, H.M.; Affan, A.; Tarhuni, N.; Raahemifar, K. Deep Learning-Based Next-Generation Waveform for Multiuser VLC Systems. Sensors 2022, 22, 2771. [CrossRef] [PubMed]
- 33. Lian, J.; Brandt-Pearce, M. Multiuser visible light communication systems using OFDMA. *J. Light. Technol.* **2020**, *38*, 6015–6023. [CrossRef]
- Asif, H.M.; Affan, A.; Tarhuni, N. Spatial Distribution Waveform Design for an Efficient Optical Communication System. In Proceedings of the 2022 4th West Asian Symposium on Optical and Millimeter-Wave Wireless Communications (WASOWC), Tabriz, Iran, 12–13 May 2022; pp. 1–5. [CrossRef]
- 35. Wang, Q.; Wang, Z.; Dai, L. Multiuser MIMO-OFDM for visible light communications. IEEE Photonics J. 2015, 7, 1–11. [CrossRef]
- 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. [CrossRef] [PubMed]
- Islam, S.M.R.; Zeng, M.; Dobre, O.A.; Kwak, K.S. Resource Allocation for Downlink NOMA Systems: Key Techniques and Open Issues. *IEEE Wirel. Commun.* 2018, 25, 40–47. [CrossRef]
- Zhang, H.; Fang, F.; Cheng, J.; Long, K.; Wang, W.; Leung, V.C. Energy-efficient resource allocation in NOMA heterogeneous networks. *IEEE Wirel. Commun.* 2018, 25, 48–53. [CrossRef]
- Liu, M.; Song, T.; Gui, G. Deep cognitive perspective: Resource allocation for NOMA-based heterogeneous IoT with imperfect SIC. *IEEE Internet Things J.* 2018, 6, 2885–2894. [CrossRef]
- Zhao, Q.; Jiang, J.; Wang, Y.; Du, J. A low complexity power allocation scheme for NOMA-based indoor VLC systems. *Opt. Commun.* 2020, 463, 125383. [CrossRef]
- 41. Ghassemlooy, Z.; Popoola, W.; Rajbhandari, S. In *Optical Wireless Communications: System and Channel Modelling with Matlab*[®]; CRC Press: Boca Raton, FL, USA, 2019.
- 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), Glasgow, UK, 11–14 May 2015; pp. 1–6.
- Wang, Q.; Zhang, R.; Yang, L.L.; Hanzo, L. Non-orthogonal multiple access: A unified perspective. *IEEE Wirel. Commun.* 2018, 25, 10–16. [CrossRef]
- 44. Breiman, L. Random forests. Mach. Learn. 2001, 45, 5–32. [CrossRef]
- 45. Biau, G.; Scornet, E. A random forest guided tour. *Test* **2016**, *25*, 197–227. [CrossRef]
- 46. Scornet, E. On the asymptotics of random forests. J. Multivar. Anal. 2016, 146, 72–83. [CrossRef]
- 47. Stitson, M.; Weston, J.; Gammerman, A.; Vovk, V.; Vapnik, V. Theory of support vector machines. Univ. Lond. 1996, 117, 188–191.
- 48. Chen, P.H.; Fan, R.E.; Lin, C.J. A study on SMO-type decomposition methods for support vector machines. *IEEE Trans. Neural Netw.* 2006, 17, 893–908. [CrossRef] [PubMed]
- Fan, R.E.; Chen, P.H.; Lin, C.J.; Joachims, T. Working set selection using second order information for training support vector machines. J. Mach. Learn. Res. 2005, 6, 1889–1918.
- 50. Cristianini, N.; Shawe-Taylor, J. An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods; Cambridge University Press: Cambridge, UK, 2000.
- 51. Fletcher, T. Support vector machines explained. In *Tutorial Paper*; UCL: London, UK , 2009; pp. 1–19.
- Tabaghi, P.; Dokmanić, I.; Vetterli, M. Kinetic Euclidean distance matrices. *IEEE Trans. Signal Process.* 2019, *68*, 452–465. [CrossRef]
 Dokmanic, I.; Parhizkar, R.; Ranieri, J.; Vetterli, M. Euclidean distance matrices: Essential theory, algorithms, and applications.
- IEEE Signal Process. Mag. 2015, 32, 12–30. [CrossRef]
- 54. Liang, Q.; Liu, M. Plugo: A vlc systematic perspective of large-scale indoor localization. arXiv 2017, arXiv:1709.06926.
- Abbas, M.; Elhamshary, M.; Rizk, H.; Torki, M.; Youssef, M. WiDeep: WiFi-based accurate and robust indoor localization system using deep learning. In Proceedings of the 2019 IEEE International Conference on Pervasive Computing and Communications (PerCom), Kyoto, Japan, 11–15 March 2019; pp. 1–10.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.