An Indoor Navigation Service for Visually Impaired People
Sarder Fakhrul Abedin, Md. Golam Rabiul Alam, Anupam Kumar Bairagi, Ashis Talukder, Choong Seon Hong
Department of Computer Science Engineering
Kyung Hee University, Korea
Email: {saab0015, robi, anupam, ashis, cshong}@khu.ac.kr

Abstract
The indoor navigation poses a diverse range of challenges due to the confined and unknown nature of the environment. Therefore, it is very difficult for the visually impaired people to navigate any indoor spaces especially for the first time. Some efforts have been made over the years to ensure a reliable and accurate indoor navigation system by using the existing technologies. However, it is also essential to maximize the efficiency of the indoor navigation system by developing a cost-effective and easy-to-manage indoor navigation as a service so that the end-user can rely on a safe and accurate indoor navigation system with low cost. In this paper we have proposed an indoor navigation service for the visually impaired people using the Bluetooth Surveillance Network (BSN). We have also modeled the navigation service using the “model free” ε greedy Q-learning algorithm. The simulation result shows the efficiency and convergence of the modeled algorithm for the proposed indoor navigation service.

1. Introduction
Navigation in a confined and unknown environment is considered as a potential challenge for the visually impaired people due to the lack of accuracy and reliability of the navigation system. Usually, in the outdoor environment, the navigation or wayfinding to reach a certain destination is largely dependent on the GPS technology. However, in the indoor environment GPS signals are unable to guarantee the desired accuracy and reliability which is required for a safe and sound navigation service for the visually impaired person. Moreover, in the indoor environment it is also crucial to mark the landmark sites (e.g. emergency exit, indoor facilities for the disabled people) and managing the map of the indoor environment efficiently and accurately.

Therefore, in this paper we have proposed the Bluetooth Surveillance Network (BSN) for the localization and indoor navigation service to the visually impaired people. In order to explore the navigation path and exploit the navigation decision in the indoor space, we have also modeled the navigation service based on the ε-greedy Q-learning algorithm for providing an efficient and reliable indoor navigation path to the visually impaired people.

2. Related Works
The researchers used the short range communication medium like RF and RFID tags [1] or integration with GPS [4], infrared camera [2] [3], WiFi signal [5] etc. In case of the infrared technology, it used the line of sight concept to detect or localize the individual user position in the indoor environment. Besides, the installation cost of such cameras for supporting multi-user becomes expensive and thus need to put more effort on the strategic placement of the cameras for ensuring accurate measurement. In case of the RF or RFID, the coverage area is so limited and requires massive deployment. So, the management of such deployment becomes a critical issue to resolve. Apart from that, the user must always carry the RFID enabled receiver device or customized white cane in order to receive and transmit signals which is hideous for an ambient navigation system. On the other hand, the WiFi signals cover comparatively large space with less deployment and installation cost. However, the WiFi signals root the cause of interference among each other which discloses another challenge for the localization of the user in an environment through WiFi signals.

3. System Overview
Fig. 1 represents the overall Bluetooth surveillance network

---

Figure 1 System Overview
and the system model for the navigation service to the visually impaired people. The Bluetooth surveillance network is composed of strategically placed Bluetooth readers (blue circle) which cover the indoor space working under a network manager (blue rectangle). Apart from that, the network manager has a graph based digital map where the nodes represent the reader locations and the landmark sites (green rectangle) (e.g. elevator, significant space) are denoted in the digital map using the Cartesian coordination system. The readers can track the position of any user with smartphone which is in discoverable mode.

A particular user presence in a specific coverage space is identified by a corresponding reader with the unique MAC address of the user owned smartphone. The user reports its presence to the reader by calculating the RSSI signal strength and the reader exchanges its ID and location coordinate (x-axis, y-axis) to the user. The user then reports the above information along with the desired destination address to the network manager. Then the manager receives the destination location from the user input (e.g. voice command to the user interface) and network manager finds the initial path between the source location and the destination. Then, the network manager sends the digital map marking the available or unavailable reader locations or nodes with the navigation directives. The user also sends the direction received from the smartphone gyroscope and the current direction of the movement to the manager. In Fig. 1, the indoor environment is changed and the initial path was A and the route is currently unavailable (grey area in the figure) due to the change in the environment. So, the alternative path B is calculated and sent by the manager. When the user reaches to the destination using path B and the user notifies the manager about the success of the path given at the updated environment. When a user detects a possible obstacle within a threshold range using ultrasonic sensor observations, it sends the signal to the user with a vibration alerts.

4. \( \varepsilon \)-greedy Q-learning algorithm

We formulate the indoor navigation problem through "model-free" \( \varepsilon \)-greedy Q-learning algorithm. In the dynamic situation, the learning update in the Q-learning is faster and reliable which is efficient in terms of a larger state space with further contribution from the exploration and exploitation or the \( \varepsilon \) value. In algorithm 1, we consider there is \( m \) number of stationary Bluetooth readers as nodes in a set \( N = \{n_1, n_2, \ldots, n_m\} \) where the Euclidean distance between the nodes are considered as links. So, the set of states is denoted as \( S = \{s_1, s_2, \ldots, s_m\} = \{n_1, n_2, \ldots, n_m\} \). The user is considered as agent and each of the Bluetooth reader establishments are considered as states where the user mobility actions from one state to another is considered as action. So, the set of action is denoted as \( A = \{a_1, a_2, \ldots, a_l\} \) where \( l \) is the possible strategies. We define a matrix \( R \) as the reward matrix where the row represents the state and column represents the action. If a state is available from the initial state there is a reward \( r \), otherwise the state is marked as unavailable or “obstacle”, the agent will be penalized. In the algorithm, Q matrix represents what the agent has learnt through experience. The rows of the Q matrix denote the current state of the agent and columns represent the possible actions leading toward the next state. The agent explores from one state to another until it converges to the goal state. Therefore, at time \( t \), for each of the action \( a_t \) in state \( s_t \) a reward \( r \) is given and the Q-value is updated as in,

\[
Q(s_t, a_t) = (1 - \alpha_t)Q(s_t, a_t) + \alpha_t[r + \beta \min Q(s_{t+1}, a_{t+1})]
\]

### Algorithm 1: \( \varepsilon \)-greedy Q-learning for the indoor navigation problem

1. Initialization
   - \( S = \{s_1, s_2, \ldots, s_m\} \) action \( A = \{a_1, a_2, \ldots, a_l\} \) for all actions \( a \), Q(s,a) arbitrary, \( r \) \( \in [0,1] \)
   - \( \alpha \in (0, 1), \beta \in (0, 1), r_{max} \in R(s,a) \)
2. While (reach episode)
3. Define initial state \( s \in S \)
4. While (state is not the goal state)
5. Find the possible action from \( A \)
6. If (a \( \in \) \( O \)) (probability for the value)
7. Exploration stage in the navigation
8. Set the action from \( A \) with the highest Q value
9. Store the cumulative reward \( \varepsilon \)
10. Count the exploitation number
11. Else // Exploration stage in the navigation
12. Take random action from \( A \)
13. Set the random action from \( A \) and calculate Q value
14. Store the cumulative reward \( r \)
15. Count the exploitation number
16. End
17. Find possible state from \( s \in S \)
18. Calculate the new Q value using equation (1) with current \( a_t, b, r \)
19. Update the value Q value for the state-action pair
20. Store the updated Q → Q(s,a)
21. Count the steps
22. End

In (1) \( Q(s_t, a_t) \) is the Q value of the current state whereas \( Q(s_{t+1}, a_{t+1}) \) is the future state. \( r \) is the expected reward at time \( t \) and \( Q(s,a) \) with the range between 0 and 1, is the learning rate \[6\]. The discount factor \( \beta \) with the range between 0 and 1. In order to train the agent, the \( \varepsilon \)-greedy Q-learning approach enables learning and exploring more states. The \( \varepsilon \)-greedy Q-learning is slightly extended from the Q-learning [7]. So, in the extended approach, at each step the random action is associated with a probability \( \varepsilon \) for the best action as per the Q-table where probability becomes 1- \( \varepsilon \). The Q matrix is updated with the Q-value in (1) when an action is performed and acts as the memory of the agent. The Q matrix thereby can be used for further navigation and the Q matrix will be updated based on the predefined \( \varepsilon \) value.
5. **Simulation**

For $\epsilon$ greedy Q-learning algorithm we have initialized the reward matrix $R$ by setting the reward value $r=50$ for a successful state change and for the unsuccessful state change we penalize the agent with $r=-50$. The number of nodes or states is 100 and when the agent reaches the goal state or destination we set the reward value $r=100$. We choose the initial state of the agent $S$ and the goal state is set to $74$. We trained the agent by 50000 episodes and set parameters $\beta = 0.8$, $\alpha = 0.90$ and $\epsilon = 0.9$. Fig. 2 represents the performance of the $\epsilon$-greedy the Q-learning algorithm for the above mentioned settings. The number of step is high at the beginning of the learning and required more episodes and then it converged after the exploration $\epsilon$. In Fig. 3 the cumulative rewards is high at the beginning because of the high number of exploration and exploitation based learning. After that, it converges as the number of exploration decreases. Based on our model and settings, the algorithm converged to the shortest path after 4175 episodes and the number of exploitation is 68897 and the number of exploration is 45129. The agent cumulative reward gain is around $4.86e+003$ at the time of convergence.

![Figure 2 $\epsilon$ greedy Q-learning performance (Steps vs Episodes)](image1)

![Figure 3 $\epsilon$ greedy Q-learning performance (Cumulative Rewards vs Episodes)](image2)

6. **Conclusion**

In this paper we have proposed the indoor navigation service for the visually impaired people using the Bluetooth Surveillance Network (BSN). We have also modeled the navigation service using the $\epsilon$ greedy Q-learning algorithm. The simulation result shows the efficiency and convergence of the modeled algorithm. In future, we will extend the research by implementing the prototype and considering diverse parameters for more reliable and accurate navigation service.

**Acknowledgement**

This research was supported by the MSIP, Korea, under the G-ITRC support program (IITP-2016-R6812-15-0001) supervised by the IITP. *Dr. CS Hong is the corresponding author.*

**References**


