Intelligent Automation & Soft Computing
DOI: 10.32604/iasc.2023.032784
Article

Tech Science Press

Reinforcement Learning-Based Handover Scheme with Neighbor Beacon Frame Transmission

Youngjun Kim[1], Taekook Kim[2], Hyungoo Choi[1], Jinwoo Park[1] and Yeunwoong Kyung[3,*]

[1]School of Electrical Engineering, Korea University, Seoul, 02841, Korea
[2]School of Computer Engineering, Pukyong National University, Busan, 48547, Korea
[3]School of Computer Engineering, Hanshin University, Osan, 18101, Korea
*Corresponding Author: Yeunwoong Kyung. Email: ywkyung@hs.ac.kr
Received: 29 May 2022; Accepted: 01 July 2022

Abstract: Mobility support to change the connection from one access point (AP) to the next (i.e., handover) becomes one of the important issues in IEEE 802.11 wireless local area networks (WLANs). During handover, the channel scanning procedure, which aims to collect neighbor AP (NAP) information on all available channels, accounts for most of the delay time. To reduce the channel scanning procedure, a neighbor beacon frame transmission scheme (N-BTS) was proposed for a seamless handover. N-BTS can provide a seamless handover by removing the channel scanning procedure. However, N-BTS always requires operating overhead even if there are few mobile stations (MSs) for the handover. Therefore, this paper proposes a reinforcement learning-based handover scheme with neighbor beacon frame transmission (MAN-BTS) to properly consider the use of N-BTS. The optimization equation is defined to maximize the expected reward to find the optimal policy and is solved using Q-learning. Simulation results show that the proposed scheme outperforms the comparison schemes in terms of the expected reward.

Keywords: Channel scanning; handover; handover delay; IEEE 802.11; WLANs; Q-learning; reinforcement learning

1 Introduction

According to the global commercialization of multimedia services, such as Netflix and YouTube, IEEE 802.11 wireless local area networks (WLANs) have proliferated due to their broadband communication capability at a reasonable cost [1]. In addition, multiple access points (APs) have been deployed in many organizations and institutions to provide mobile multimedia services to users anytime and anywhere [2,3].

Originally, WLAN was introduced for static (i.e., not mobile) users within a limited coverage [4]. However, as users become mobile, mobility support to change the connection from the current AP (CAP) to the next (i.e., handover) becomes an important issue [5]. Since handover includes searching for neighbor APs (NAPs) in different channels (i.e., channel scanning), authentication to the target AP, and reassociation to that AP, handover is time-consuming; consequently, service quality deteriorates, especially for delay-sensitive services. During handover, the channel scanning procedure, which aims to collect NAP information in all of the available channels, accounts for most of the delay time [5,6]. Therefore, it is necessary to find an efficient way to reduce the channel scanning procedure for fast handover.

This situation leads to the need for AP coordination or centralized AP control, usually known as the enterprise WLAN [2,6,7]. In studies on enterprise WLANs, there have been efforts to reduce the channel scanning procedure, which aims to allow mobile stations (MSs) to recognize the channel information of NAPs. These works can be categorized into (1) MS-based approaches [8,9] and (2) network-based approaches [10,11]. Even though virtual AP management has recently been introduced for mobility support in enterprise WLANs, virtual AP management aims to remove authentication and reassociation (i.e., not to improve the channel scanning procedure) during handover through virtual AP migration between APs [12–14].

In the MS-based approach, an MS can find NAPs by obtaining their channel information through the request/response (i.e., neighbor report defined by 802.11k or basic service set transition management provided by 802.11v) with the controller [8,15] or channel switching by itself [9]. If the controller exists, the former is more efficient because of the channel switching overhead at each MS. On the other hand, in the network-based approach, APs share the channel information with MSs through the channels of NAPs by using only one interface (i.e., used for both data communications and sending channel information) [10] or an additional interface (i.e., dedicated for channel scanning purposes) [11]. Using only one interface is more efficient in terms of the hardware cost and resource utilization.

In this paper, we propose a reinforcement learning-based handover scheme with neighbor beacon frame transmission (MAN-BTS), where the controller in enterprise WLANs determines whether to use the MS-based approach or network-based approach at each time epoch. If the controller decides to use the MS-based approach, MSs for handover try to obtain the NAP information by themselves through the request/response with the controller. In terms of the handover delay, after MSs try to obtain the NAP information, an active scanning procedure for the NAP channels is needed; this procedure can be time-consuming. Additionally, if a network-based approach is used, when there are few MSs for handover, periodical transmissions of neighbor-beacon frames (NBFs) are inefficient in terms of the handover preparation cost. Therefore, the optimization equation is defined to maximize the expected reward to find the optimal policy and is solved using Q-learning (QL). Simulation results show that MAN-BTS outperforms the comparison schemes in terms of the expected reward.

The contribution of this paper is threefold: (1) to the best of our knowledge, this is the first study where network-based and MS-based handover approaches are adaptively considered to find an optimal handover policy; (2) to determine the WLAN handover policy, we considered not only the distribution of MSs but also the presence of delay-sensitive MSs; and (3) by means of simulations, the performance of MAN-BTS is evaluated under various environments; these evaluations can provide a valuable design for mobility support in enterprise WLANs.

The remainder of this paper is organized as follows. After related work is presented in Section 2, MAN-BTS is described in Section 3. The simulation results and concluding remarks are discussed in Sections 4 and 5, respectively.

2 Related Work

Studies on handovers to improve the performance of channel scanning procedures have been reported in the literature [8–11]. These studies can be classified into two categories, namely, MS-based approaches [8,9] and network-based approaches [10,11], depending on the main agent performing the channel scanning procedure.

Using an MS-based approach, Zhang et al. [8] proposed a neighbor list proactive (NLP) handover scheme based on the neighbor list, which is managed by a controller. The MS requests and receives an NAP list from the controller. Due to the list, the channel scanning procedure can be efficient in reducing overhead because only several attempts by MSs are needed to obtain the information of NAPs for handover. However, even though the number of scanning channels is reduced, MSs' active scanning procedure is still required for MSs' NAP channels. Ramani et al. [9] proposed a SyncScan, where the clocks of all APs are synchronized. In this research, the time to transmit beacon frames is manually configured and set according to the channel. AP transmit beacon frames at a fixed time by using a predetermined channel. The MS knows at what time and to what channel beacons come so that the MS can receive beacon frames by changing the channel according to the time the beacon frames are about to arrive. This allows beacons to be received by changing channels on time without having to perform passive scanning or active scanning for beacon reception. However, the power consumption to keep receiving the beacon frames and the overhead of continuously changing the channel become too large.

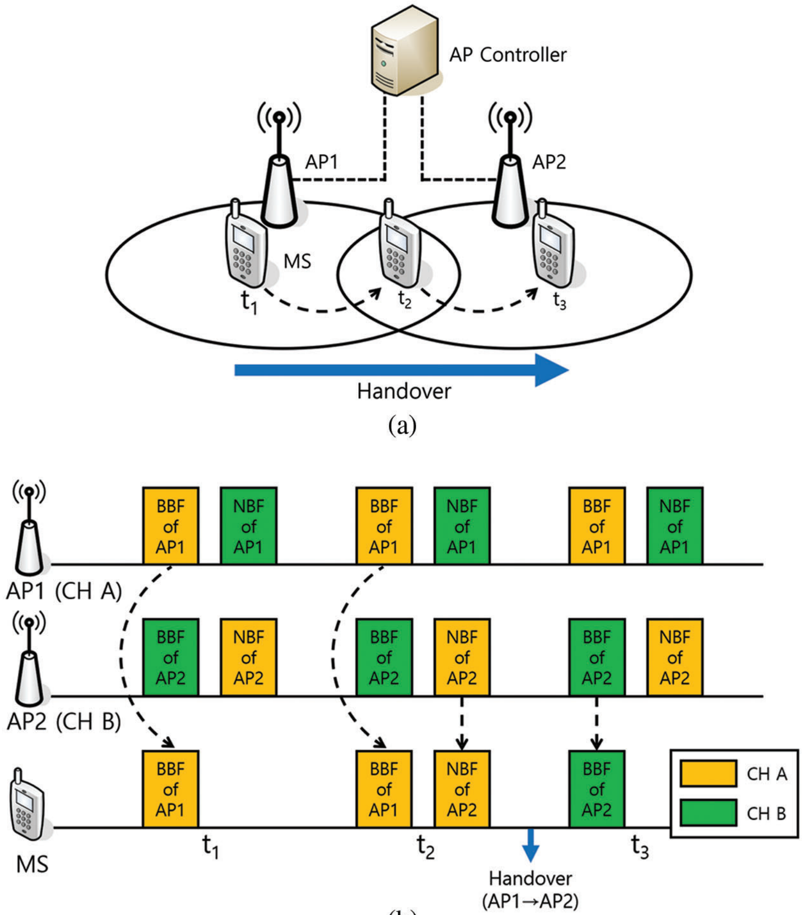
In a network-based approach, Kim et al. [10] proposed a neighbor beacon frame transmission scheme (N-BTS). Each AP periodically sends beacon frames, including the operation channel information, by using the NAP channel according to a predetermined order. In this way, the target NAP for handover can be determined from the MS's view, and the channel scanning procedure can thus be eliminated. However, periodical multiple beacon frame transmission is always needed and can thus be overhead (i.e., cost) when there are few MSs for handover. Jeong et al. [11] used dual network interfaces, the primary one for data communication with the MS and the secondary one dedicated to transmitting beacon frames by using channels of NAPs. Through this, instead of channel scanning, MS can receive beacons sent by secondary network interfaces of NAPs to the operating channel of the MS to obtain information from NAPs. Although the handover delay can be greatly reduced by omitting the channel scanning step, using two network interfaces is quite price inefficient. Two network interfaces have an opportunity cost of more than double the system throughput.

3 MAN-BTS

In this section, we describe the network architecture and the operation of MAN-BTS. In addition, a method to find the optimal policy for MAN-BTS was introduced. Compared with our previous work, N-BTS [10], MAN-BTS supports both MS-based and network-based approaches, and the AP controller can determine the optimal handover policy by considering the network status. MAN-BTS provides seamless handover, reduces the handover overhead and maximizes the expected reward in IEEE 802.11 WLANs. In MAN-BTS, each AP periodically transmits beacon frames through the channels of NAPs in an order received from the AP controller, which contains the operating channel information of the AP. This approach allows MSs to receive the beacon frames of NAPs, and therefore MSs can determine the appropriate AP that they are heading to without the channel scanning procedure. In addition, MAN-BTS can be adaptively operated to consider the network status. Performance evaluation results demonstrate that MAN-BTS can maximize the expected reward by using the optimal policy.

3.1 Network Architecture of MAN-BTS

Fig. 1a shows the network architecture of MAN-BTS. In this network, it is assumed that all APs are connected to the AP controller and create an extended service set (ESS), which can be considered a single basic service set from the MSs' view. In Fig. 1a, AP1 and AP2 are connected to the AP controller, which coordinates the channels used by APs and synchronizes the timing for beacon transmissions of APs in the ESS. In addition, since we assume the enterprise network [16,17], the AP controller can have NAP lists of APs. In addition, APs are assumed to be strategically deployed to minimize overlapping communication areas between APs and maximize the overall coverage of extended service sets. Channel reuse allows networks to be configured using three channels (either 2.4 or 5.8 GHz band). In other words, when using MAN-BTS, a maximum of two NBFs can be transmitted when the network is configured using three channels. The MS in Fig. 1a is connected to AP1 from t_1 to t_2 and moves in the direction of AP2. After t_2, handover is performed to connect to AP2, and the MS is connected to AP2 at t_3.



(a)

(b)

Figure 1: MAN-BTS: (a) Network architecture, (b) Timing diagram of the neighbor beacon frame transmission

3.2 Network-based Operation of MAN-BTS

In MAN-BTS, the CAP communicates with the MS on a single operational channel but transmits an additional NBF over the channel of the NAP during the beacon frame transmission period. First, the beacon frame (i.e., the basic beacon frame (BBF) is delivered using the AP channel, and the NBF is transmitted using the operating channels of the NAPs. The AP controller manages and determines the order and timing in which the AP transmits one BBF and an additional NBF. Although the role of the BBF is the same as that of conventional beacon frames, the NBF is intended to announce existence of the AP to MSs connected to other NAPs and to inform the operating channel. When the MS receives beacon frames from the service AP and the NAP, the MS can compare the received signal strength indicator (RSSI) of the received beacon frames without the channel scanning procedure and determine the best AP for handover according to the real-time RSSI.

Fig. 1b illustrates the timing diagram of the neighbor beacon frame transmission; the diagram shows the reception status of the beacon frames from the perspective of the MS at times t_1, t_2 and t_3 as the MS moves from AP1 to AP2. At time t_1, the MS is connected to AP1 and set to Channel A (CH A). Since AP1 is the only adjacent AP of the MS, the MS receives the BBF of AP1. When the MS is located between AP1 and AP2 at time $t_{2,}$ the MS receives not only the BBF of the currently connected AP1 but also the NBF of AP2 through CH A. When two adjacent APs periodically send beacon frames, the MS receives the beacon frames and decides to start the handover. In other words, the MS can compare the RSSIs of beacon frames from AP1 and AP2 and decide when and to which AP the MS needs to perform the handover [18–21]. In addition, other handover prediction methods can be applied by obtaining information through the NBF reception [22,23]. As an example, the MS performs the handover from AP1 to AP2 after t_2. At t_3, the MS is located in the communication area of AP2, communicates with AP2 through Channel B, and receives only the BBFs transmitted by AP2. Active scanning by the MS is not required during the handover, and only the APs need to transmit additional NBFs. Therefore, the channel scanning procedure is omitted, and thus, the handover delay time can be significantly reduced. In addition, since the MS can constantly measure and compare the RSSIs of the beacon frames from the NAP and the CAP, it is possible to determine the handover more accurately than determining the handover based on the RSSI from the CAP. Even though MAN-BTS requires modification, this modification is possible with a simple software upgrade because only one handover decision rule needs to be added.

If some beacon frames are sent to the same channel at the same time, they cannot be received due to a collision. Therefore, time synchronization to prevent beacon frame collisions between the APs is essential. To resolve this issue, the AP controller synchronizes the timing of the BBF and the NBF and notifies the APs of the timing. However, even if the AP controller manages the beacon frame transmission timing of the APs, the wireless media access of the WLAN is a competitive approach, so the AP is not guaranteed to transmit beacon frames at a predetermined time. Thus, a margin is placed between the BBF and the NBF to ensure the transmission timing of the NBF. A network allocation vector (NAV) is established in the BBF to prevent MSs from using the channel until the beacon frame transmission is completed.

3.3 MS-based Operation of MAN-BTS

The network-based operation of MAN-BTS is effective when there are many MSs close to the handover. On the other hand, the MS-based operation of MAN-BTS is introduced in this part, which is advantageous in situations where there are few MSs close to the handover. In MAN-BTS, each AP monitors the RSSI of the MSs and reports an AP controller if the RSSI is below the handover preparation threshold (th_{HP}). The AP controller knows the number of MSs close to the handover through the reports from the APs. If the RSSI of the MS falls below the handover decision threshold (th_{HD}), the AP disconnects with that MS. During the disconnection procedure, the MS requests the channel information of the NAPs to the AP, and the AP responds to the MS after obtaining the information of the NAP through the AP controller. Then, the MS starts the selective channel scanning procedure to find new APs. In this way, handover becomes faster than performing a channel scanning procedure for all channels, and lower overhead occurs when a few MSs perform the handover. The handover preparation area (HPA) is a region with an RSSI of th_{HP} or less and th_{HD} or more, as shown in Fig. 2. In this way, compared to network-based operations, the MS-based method requires an NAP list request/response only when a handover occurs. Therefore, the MS-based operation of MAN-BTS is suitable when the number of MSs and the number of delay-sensitive MSs located in the HPA are small.

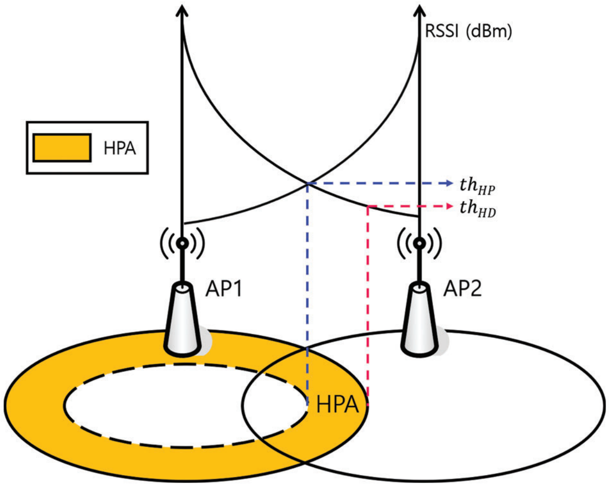


Figure 2: The HPA and the HPA criteria for AP1

3.4 Optimal Policy for MAN-BTS

As we mentioned above, although the network-based operation of MAN-BTS supports the handover decision of the MS and can significantly reduce the handover latency, there is a burden of periodic transmissions of the additional beacons, NBFs (i.e., the AP occupies the wireless medium during the channel switching time and NBF transmission time). That is, the network-based operation of MAN-BTS has a trade-off between the gain of increasing the performance of the handover and the overhead for transmitting additional NBFs. Consequently, the network-based operation of MAN-BTS can benefit greatly in the case of frequent handovers. However, when there is no MS for the handover, only overhead can occur. Therefore, MAN-BTS, which can determine the optimal policy to use MAN-BTS considering the real-time distribution of the MSs, is proposed. The AP controller determines the MAN-BTS usage policy based on the handover proximity of the real-time MS distribution collected from the APs. The AP controller makes policy decisions based on the number of MSs and the number of delay-sensitive MSs in the HPA. An optimization equation is defined that aims to maximize the expected reward, and then a QL-based MAN-BTS utilization decision mechanism is proposed.

3.4.1 Q Learning

Machine learning has been utilized to make predictions and/or decisions [24,25]. Among the categories of machine learning, reinforcement learning (RL), also known as enhancement learning, is inspired by a behaviorist sensibility and control psychology. The main idea is for an agent to perform actions in the environment and maximize the expected reward through learning experiences. QL is the most widely used RL technique and has been extensively applied in wireless networks [26–28].

QL uses the state-action value function Q. It is assumed that we do not know the state that we reach when we select each action from a given state. However, the Q-value function allows us to know the value of the action selection. Consequently, we can choose the action with the highest value. The QL algorithm finds the optimal Q-value iteratively by updating the Q-value as follows:

$$Q(S,A) \leftarrow Q(S,A) + \alpha \cdot (R + \gamma \cdot \max Q(S',A') - Q(S,A)) \tag{1}$$

where α $(0 \leq \alpha \leq 1)$ denotes the learning rate, which determines to what extent the newly learned information affects the old information. The larger α is, the higher the degree to which the new information is valued. $\gamma(0 \leq \gamma \leq 1)$ indicates the discount factor. Before beginning QL, the Q-value table is initialized to zeroz, and at each subsequent iteration, the Q-value is updated until the final Q-value table converges. In the convergence Q-value table, the action corresponding to the maximum Q-value of each state is each state's optimal action obtained by the learning. However, if the Q-value is updated through the greedy policy in the above iteration, it can be difficult to explore sufficiently because the agent repeatedly selects only one action with the high value in a specific state. Therefore, the agent should not always greedily select the action and should sometimes select other actions that are not of high value to ensure the exploration. In this paper, we use the ε-greedy exploration-exploitation method for proper exploration. At each decision epoch, the agent randomly selects the action with probability ε. Then, the greedy action is exploited based on the previously learned Q-value with probability $1 - \varepsilon$. The formula is as follows:

$$\pi(A|S) = \begin{cases} 1 - \varepsilon \text{ if } a^* = \arg\max\limits_{A} Q(S, A) \\ \qquad\qquad \varepsilon \text{ otherwise} \end{cases} \tag{2}$$

This greedy method guarantees some exploration. In addition, in this paper, we use the decaying ε-greedy method, in which the ε value is initially set as a high value and then reduced by δ after each iteration. At first, the agent knows nothing about the environment. Then, the agent focuses on obtaining enough information about the environment by selecting various actions. After learning to some extent, the agent focuses on making the best choice based on the information already obtained. Algorithm 1 shows steps to select the optimal Q-value by applying the decaying ε-greedy method. When the Q-value table converges, the agent can find the optimal policy π_{opt} that maximizes the expected long-term discount reward $Q(S, A)$; i.e., $\pi_{opt} = \arg\max\limits_{A} Q(S, A)$.

Algorithm 1: Steps for the optimal Q-value table

```
Step 1: For each S and A, initialize $Q(S, A)$ to zero. Specify $\varepsilon > 0$;
Step 2: Observe the current state $S$;
Step 3: Select an action, and execute $A$ depending on Eq. (2);
Step 4: Receive immediate returns $R$;
Step 5: Observe the new state $S'$;
Step 6: Update the Q-value table according to Eq. (1);
Step 7: The Q-value table converges and ends; otherwise, $\varepsilon = \varepsilon - \delta$, and repeat Steps 2~6.
```

3.4.2 State Definition

To efficiently utilize MAN-BTS, it is important to first properly measure the situation of the network. State space $\boldsymbol{S}$ is defined as follows:

$$\boldsymbol{S} = \boldsymbol{N} \times \boldsymbol{U} \tag{3}$$

where $\boldsymbol{N}$ and $\boldsymbol{U}$ denote the handover preparation phase of the MSs and the handover preparation phase of the delay-sensitive MSs, respectively.

When the total number of MSs in the network is M, $\boldsymbol{N}$ and $\boldsymbol{U}$ are represented by

$$N = \{0, 1, 2, \ldots, M\} \tag{4}$$

$$U = \{0, 1, 2, \ldots, M\} \tag{5}$$

where $n(\in \boldsymbol{N})$denotes the situation when the number of MSs located in HPA is n and u $(\in \boldsymbol{U})$denotes the situation when the number of delay-sensitive MSs located in HPA is u.

3.4.3 *Action Definition*

In MAN-BTS, the AP controller selects whether to execute the network-based MAN-BTS according to the current state information. The action $\boldsymbol{A}$ can be described as

$$\boldsymbol{A} = \{0, 1\} \tag{6}$$

where $a(\in \boldsymbol{A})$represents whether the AP controller commands a network-based or MS-based operation of MAN-BTS execution. If $a = 0$, MS initiates a channel scanning procedure for a handover when RSSI falls below th_{HD}. However, in the case of $a = 1$, MS compares the RSSI of beacon frames from the CAP and the NAP and performs the handover when the RSSI of the NAP is three times higher than that of the CAP.

3.4.4 *Reward and Cost Functions*

The reward and cost functions depend on state S and action A. The reward function includes the reward for the handover, and the cost function is related to the cost needed to prepare the handover; this cost includes the cost of the beacon transmissions and the cost of the channel scanning time. The total reward function, $R(S, A)$, is defined as

$$R(S, A) = w \cdot f(S, A) - (1 - w) \cdot g(S, A) \tag{7}$$

where $f(S, A)$ denotes the reward function for the handover and $g(S, A)$ denotes the cost function due to the handover preparation. In addition, $w(0 \le w \le 1)$ denotes the weight factor to adjust the significance of $f(S, A)$ and $g(S, A)$.

The reward function according to the action, $f(S, A)$, is defined as

$$f(S, A) = \begin{cases} r_{DT} \cdot H_{DT}, & \text{if } a = 0 \\ r_{DT} \cdot H_{DT} + r_{DS} \cdot H_{DS}, & \text{if } a = 1 \end{cases} \tag{8}$$

where r_{DT} and r_{DS} indicate the reward for the handover of the delay-tolerant MS and the delay-sensitive MS, respectively; and where H_{DT} and H_{DS} denote the number of handovers for the delay-tolerant MSs and the number of handovers for the delay-sensitive MSs, respectively. Since the network-based operation of MAN-BTS can satisfy the delay requirement of the delay-sensitive MS, r_{DS} can be obtained when $a = 1$.

However, the overhead due to using MAN-BTS should be considered. Consequently, the cost function, $g(S, A)$, is defined as

$$g(S, A) = \begin{cases} c_0 \cdot (H_{DT} + H_{DS}), & \text{if } a = 0 \\ c_1, & \text{if } a = 1 \end{cases} \tag{9}$$

where c_0 refers to the cost of the channel scanning procedure by MS for a handover; this cost increases in proportion to the number of MSs performing handovers. In the case of using the network-based operation of MAN-BTS (i.e., $a = 1$), a channel scanning procedure is not needed. However, there is a periodical overhead for the AP to transmit the NBF. In this case, a constant cost c_1 occurs regardless of the number of handovers.

4 Evaluation Result

For the performance evaluation, we compared the proposed MAN-BTS with (1) N-BTS [10], where MS can obtain the NBFs from the NAPs; and (2) NLP [8], where MS can obtain the channel information of the NAPs through an AP controller [8]. We assume that 7 APs are placed in hexagonal cell deployment and the MS moves according to the random-walk model [29]. Tab. 1 includes the parameters used in the performance evaluation.

Table 1: Default parameter setting

Parameters	M	w	q	r_{DT}	r_{DS}	c_0	c_1
Value	10	0.7	0.2	2	5	−1	−5

4.1 Effect of w

The effect of the weighted factor w on balancing the reward functions is shown in Fig. 3. As shown in Fig. 3, the expected reward value of MAN-BTS is always the highest. When the w value is from 0.1 to 0.4 (i.e., a higher weight is given to the cost rather than the reward), MAN-BTS has almost the same expected reward value as NLP. That is, MAN-BTS operates in almost NLP mode. Additionally, when the value of w is between 0.8 and 0.9, the expected rewards of MAN-BTS and N-BTS are close to each other. When w is between 0.5 and 0.7, the expected reward value of MAN-BTS is higher than that of the others by an appropriate policy decision between the NLP and N-BTS modes according to the state.

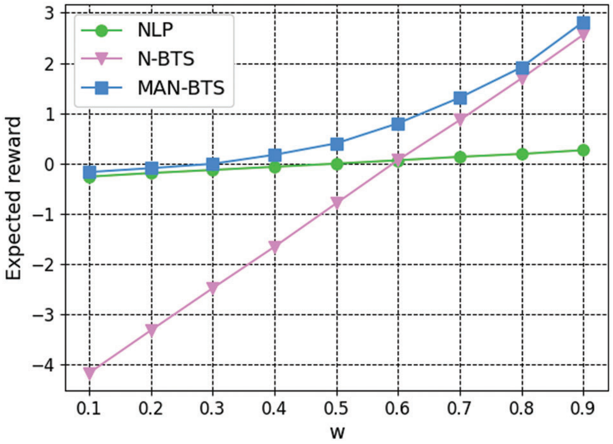


Figure 3: Expected reward according to weight factor w

4.2 Effect of q

Fig. 4 shows the expected reward as a function of waiting probability q. q denotes the probability that the MS leaves from a place during the decision epoch. Consequently, when q is high, the MS moves actively. However, when q is low, the MS corresponds to a static case in which the MS remains stationary. MAN-BTS has the highest expected reward irrespective of the change in q. When q is greater than 0.6, MAN-BTS has almost the same value as NLP. The performance of NLP and N-BTS crosses when q is between 0.2 and 0.3, and when q is greater than 0.3, the expected reward value of NLP becomes higher.

4.3 Effect of M

M denotes the total number of MSs in the access network. As M increases, the probability of handover increases. Fig. 5 shows the expected reward value when M is from 5 to 15. As M increases, the expected reward of N-BTS increases and that of NLP decreases. The expected reward value of MAN-BTS is always the highest in the entire domain. Specifically, as M increases, handover occurs more frequently, and in such a situation, N-BTS is more efficient. However, NLP is more efficient when handover rarely occurs (i.e., M is small). In contrast, MAN-BTS always shows an optimal result because MAN-BTS adaptively operates according to M.

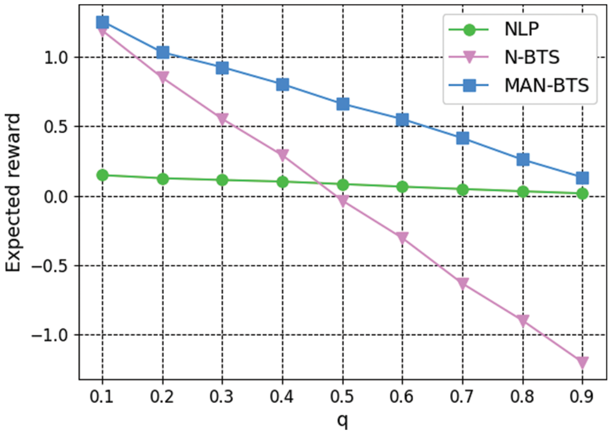


Figure 4: Effect of waiting probability q on the expected reward

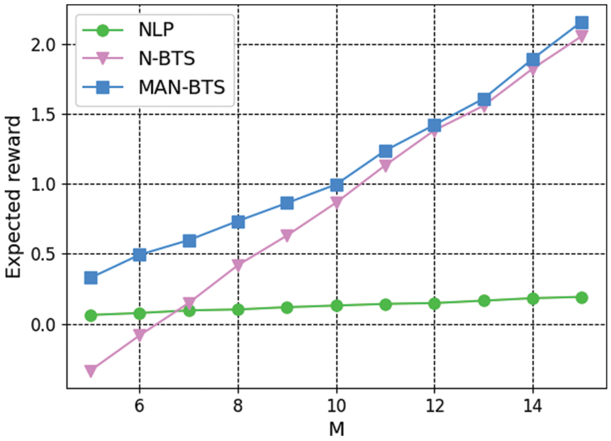


Figure 5: Effect of the total number of MSs on the expected reward

5 Conclusion

Through the network-based operation of MAN-BTS, MSs can receive the NBFs of NAPs, thereby enabling MSs to check the surrounding AP information in real time and make appropriate handover decisions. In this paper, we propose MAN-BTS to make efficient use of MAN-BTS according to the network situation. MAN-BTS can find the optimal policy to maximize the expected reward by using Q-learning. By considering the higher reward on delay-sensitive MSs, we propose a more practical method instead of a uniform handover policy. The evaluation results demonstrate that compared to NLP and N-BTS, MAN-BTS with the optimal policy can always achieve the highest expected reward. In our future work, we will extend MAN-BTS to consider the routing path to minimize the packet loss and service delay in response to handover under the control of the AP controller.

Acknowledgement: This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea Government (MSIT) (No.2020R1G1A1100493).

Funding Statement: This work was funded by the Korea Government (MSIT).

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

[1] S. Manzoor, Z. Chen, Y. Gao, X. Hei and W. Cheng, “Towards QoS-aware load balancing for high density software defined Wi-Fi networks,” *IEEE Access*, vol. 8, pp. 117623–117638, 2020.

[2] M. Peng, G. He, L. Wang and C. Kai, “AP selection scheme based on achievable throughputs in SDN-enabled WLANs,” *IEEE Access*, vol. 7, pp. 4763–4772, 2019.

[3] Q. M. Kharma, A. H. Hussein, F. M. Taweel, M. M. Abualhaj and Q. Y. Shambour, “Investigation of techniques for voip frame aggregation over a-mpdu 802.11n,” *Intelligent Automation & Soft Computing*, vol. 31, no. 2, pp. 869–883, 2022.

[4] A. Hills, “Large-scale wireless LAN design,” *IEEE Communications Magazine*, vol. 39, no. 11, pp. 98–107, 2001.

[5] I. Purushothaman and S. Roy, “FastScan: A handoff scheme for voice over IEEE 802.11 WLANs,” *Wireless Networks*, vol. 16, no. 7, pp. 2049–2063, 2010.

[6] J. Q. Filho, N. Cunha, R. Lima, E. Anjos and F. Matos, “A software defined wireless networking approach for managing handoff in IEEE 802.11 networks,” *Wireless Communications and Mobile Computing*, vol. 2018, no. 9246824, pp. 1–12, 2018.

[7] A. Zubow, S. Zehl and A. Wolisz, “BIGAP–Seamless handover in high performance enterprise IEEE 802.11 networks,” in *Proc. NOMS*, Istanbul, Turkey, pp. 445–453, 2016.

[8] H. Zhang, Z. Lu, X. Wen and Z. Hu, “QoE-based reduction of handover delay for multimedia application in IEEE 802.11 networks,” *IEEE Communications Letters*, vol. 19, no. 11, pp. 1873–1876, 2015.

[9] I. Ramani and S. Savage, “SyncScan: Practical fast handoff for 802.11 infrastructure networks,” in *Proc. INFOCOM*, Miami, FL, USA, 1, pp. 675–684, 2005.

[10] Y. Kim, H. Choi, K. Hong, M. Joo and J. Park, “Fast handoff by multi-beacon listening in IEEE 802.11 WLAN networks,” in *Proc. ICUFN*, Milan, Italy, pp. 806–808, 2017.

[11] J. Jeong, Y. D. Park and Y. Suh, “An efficient channel scanning scheme with dual-interfaces for seamless handoff in IEEE 802.11 WLANs,” *IEEE Communications Letters*, vol. 22, no. 1, pp. 169–172, 2018.

[12] L. Sequeira, J. L. de la Cruz, J. Ruiz-Mas, J. Saldana, J. Fernandez-Navajas *et al.,* “Building an SDN enterprise WLAN based on virtual APs,” *IEEE Communications Letters*, vol. 21, no. 2, pp. 374–377, 2017.

[13] J. L. Vieira and D. Passos, “An SDN-based access point virtualization solution for multichannel IEEE 802.11 networks,” in *Proc. NoF*, Rome, Italy, pp. 122–125, 2019.

[14] E. Zeljković, N. Slamnik-Kriještorac, S. Latré and J. M. Marquez-Barja, “ABRAHAM: Machine learning backed proactive handover algorithm using SDN,” *IEEE Transactions on Network and Service Management*, vol. 16, no. 4, pp. 1522–1536, 2019.

[15] M. I. Sanchez and A. Boukerche, “On IEEE 802.11K/R/V amendments: Do they have a real impact?,” *IEEE Wireless Communications*, vol. 23, no. 1, pp. 48–55, 2016.

[16] S. Latif, S. Akraam, A. J. Malik, A. A. Abbasi, M. Habib *et al.,* “Improved channel allocation scheme for cognitive radio networks,” *Intelligent Automation & Soft Computing*, vol. 27, no. 1, pp. 103–114, 2021.

[17] B. Dezfouli, V. Esmaeelzadeh, J. Sheth and M. Radi, “A review of software-defined WLANs: Architectures and central control mechanisms,” *IEEE Communications Surveys & Tutorials*, vol. 21, no. 1, pp. 431–463, 2019.

[18] V. Pichaimani and K. R. Manjula, “A machine-learning framework to improve wi-fi based indoor positioning,” *Intelligent Automation & Soft Computing*, vol. 33, no. 1, pp. 383–397, 2022.

[19] N. Singh, S. Choe and R. Punmiya, “Machine learning based indoor localization using Wi-Fi RSSI fingerprints: An overview,” *IEEE Access*, vol. 9, pp. 127150–127174, 2021.

[20] D. D. Nguyen and M. Thuy Le, “Enhanced indoor localization based BLE using gaussian process regression and improved weighted KNN,” *IEEE Access*, vol. 9, pp. 143795–143806, 2021.

[21] S. Zhao, F. Wang, Y. Ning, Y. Xiao and D. Zhang, “Vertical handoff decision algorithm combined improved entropy weighting with GRA for heterogeneous wireless networks,” *KSII Transactions on Internet and Information Systems*, vol. 14, no. 11, pp. 4611–4624, 2020.

[22] L. Huang, L. Lu and W. Hua, “A survey on next-cell prediction in cellular networks: Schemes and applications,” *IEEE Access*, vol. 8, pp. 201468–201485, 2020.

[23] J. Chen, J. Li, M. Ahmed, J. Pang, M. Lu *et al.,* “Next location prediction with a graph convolutional network based on a Seq2seq framework,” *KSII Transactions on Internet and Information Systems*, vol. 14, no. 5, pp. 1909–1928, 2020.

[24] H. Sun and R. Grishman, “Employing lexicalized dependency paths for active learning of relation extraction,” *Intelligent Automation & Soft Computing*, vol. 34, no. 3, pp. 1415–1423, 2022.

[25] H. Sun and R. Grishman, “Lexicalized dependency paths based supervised learning for relation extraction,” *Computer Systems Science and Engineering*, vol. 43, no. 3, pp. 861–870, 2022.

[26] C. Ke and L. Astuti, “Applying deep reinforcement learning to improve throughput and reduce collision rate in IEEE 802.11 networks,” *KSII Transactions on Internet and Information Systems*, vol. 16, no. 1, pp. 334–349, 2022.

[27] H. Ko, S. Pack and V. C. M. Leung, “Mobility-aware vehicle-to-grid control algorithm in microgrids,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 19, no. 7, pp. 2165–2174, 2018.

[28] Y. Chen, J. Zhao, Q. Zhu and Y. Li, “Research on unlicensed spectrum access mechanism based on reinforcement learning in LAA/WLAN coexisting network,” *Wireless Networks*, vol. 26, no. 3, pp. 1643–1651, 2020.

[29] I. F. Akyildiz and W. Wang, “A dynamic location management scheme for next-generation multitier PCS systems,” *IEEE Transactions on Wireless Communications*, vol. 1, no. 1, pp. 178–189, 2002.