

Received February 6, 2018, accepted March 29, 2018, date of publication March 15, 2018, date of current version April 23, 2018. *Digital Object Identifier* 10.1109/ACCESS.2018.2815990

Hybrid Slotted-CSMA/CA-TDMA for Efficient Massive Registration of IoT Devices

NURULLAH SHAHIN, RASHID ALI, AND YOUNG-TAK KIM[®], (Member, IEEE)

Department of Information and Communication Engineering, Graduate School, Yeungnam University, Gyeongsan 38541, South Korea Corresponding author: Young-Tak Kim (ytkim@yu.ac.kr)

This work was supported in part by the Ministry of Science and ICT, South Korea, under the Information Technology Research Center support program supervised by the Institute for Information and Communications Technology Promotion under Grant IITP-2017-2016-0-00313, in part by the Yeungnam University 2015 Research Grant and the Fellowship Grants for Higher Education and Research in Information and Communication Technology, and in part by the Ministry of Post, Telecommunication and Information Technology, Bangladesh, in 2015, under Grant 3-2801-0001-5963.

ABSTRACT Recently, the Wi-Fi alliance announced a new Wi-Fi standard known as IEEE 802.11ah (or Wi-Fi HaLow) to efficiently support Internet of Things (IoT) applications. However, the existing registration method under IEEE 802.11ah, based on carrier-sense multiple access with collision avoidance (CSMA/CA), was analyzed as not efficient enough for registration of large-scale machine-to-machine (M2M) communications where a massive number of devices try to access a single, centralized access point (AP). In this paper, we propose a hybrid slotted-CSMA/CA-time-division multiple access (TDMA) (HSCT) medium access control (MAC) protocol for efficient massive registration of IoT devices (up to 8000) in M2M networks. We focus on situations, where a large number of M2M devices simultaneously try to register at a single, centralized AP. In the proposed HSCT, contention-based slotted-CSMA/CA allows devices to send an authentication request via randomly selected backoff slots, whereas contention-free TDMA permits those devices to send/receive the subsequent association request/association response via an individually allocated TDMA slot. In addition, a centralized authentication control (CAC)-based mechanism with modified algorithms for optimal selection of CAC parameters and the slotted fixed-window CSMA protocol with Sift geometric probability distribution are used to mitigate severe contention between massive registrations upon network (re-)initialization from an AP reboot. This paper also analyzes the performance of the proposed scheme and determines the optimal configuration to enhance registration performance. Simulation results demonstrate that the proposed HSCT MAC protocol achieves substantial improvement, compared with the contention-free transmission, a combined authentication/association scheme, and the conventional IEEE 802.11ah with CSMA/CA.

INDEX TERMS M2M networks, IEEE 802.11ah, Internet of Things (IoT), authentication, association, hybrid CSMA/CA-TDMA.

I. INTRODUCTION

Machine-to-machine (M2M) communications is an essential part of the emerging Internet of Things (IoT), which exchanges information among autonomous sensors/actuators without human interaction [1]. The future generations of wireless IoT devices are expected to be intelligent and more efficient, with interconnections to the global Internet. The deployment of smart devices serving various IoT applications is estimated to grow to over 30 billion globally by 2020 [2]. The IEEE 802.11ah Task Group [3] is working on a draft amendment for standardization that addresses efficient M2M network support for large numbers of devices, long transmission ranges, short and infrequent data transmissions, and very low power consumption [4]. Smart and efficient management of massive registrations is one of the key requirements needed to build scalable, flexible, and dynamic networks for IoT applications.

In order to handle large numbers of devices in M2M communications, an IEEE 802.11ah wireless local area network (WLAN) must support up to 8000 devices connected to a single access point (AP) with a transmission range of up to 1 km [4]. A registration procedure must be completed before exchanging sensor/actuator data from/to devices. A four-way handshake mechanism is required to complete the registration process, which includes an authentication request (AuthReq), authentication response (AuthResp), an association request

(AssocReq), and an association response (AssocResp). All devices in the network send AuthRegs and AssocRegs, and the AP responds with AuthResps and AssocResps. The devices obtain an association identification (AID) and get permission to exchange application data after a successful registration procedure [3]. Currently, in an IEEE 802.11ah WLAN, carrier-sense multiple access with collision avoidance (CSMA/CA) is used to exchange the request/response messages of the registration process. Even though CSMA/CA is a popular contention-based random access protocol with high flexibility, scalability, and robustness, the congestion level gradually increases as the network grows. Therefore, CSMA/CA is not efficient enough when huge numbers of M2M devices try to access a single centralized AP all at once [5]. Moreover, due to the four-way handshake in the registration process, every IoT device must access the channel twice (i.e., once for the AuthReq and again for the AssocReq), so if 8000 devices are in the M2M network, a total of 16,000 AuthReq and AssocReq messages are sent to the AP [5]. Therefore, massive numbers of CSMA/CA-based accesses produce severe collisions that result in a long time to complete the registration procedures.

Unlike CSMA, time division multiple access (TDMA) is a collision-free access scheme that avoids competition for channel access [6]. The transmission time is divided into slots, where each slot is allocated to a device via appropriate scheduling, and each device attempts to transmit only during its assigned TDMA slot (T-slot). The major advantage of TDMA is higher channel utilization, because there are no collisions during channel access. The coordinator (i.e., the AP) first checks the availability of T-slots, assigns the available T-slots, and informs the nodes of their allocated T-slot. However, before the registration procedure, the AP does not have any of the information required for proper TDMA scheduling in a real environment.

At any time, the network may have to restart or re-initialize for various reasons, such as an AP reboot, a system crash, a power failure, and so on. Once the AP restarts, devices simultaneously try to reconnect, and the whole registration process takes a long time to complete for up to 8000 devices. There are several approaches to mitigating severe contention during device registrations. The IEEE 802.11ah standard introduced a centralized authentication control (CAC) method that limits the number of devices accessing the communications channel to send AuthReq and AssocReq messages [5]. This method offers several parameters to achieve the optimal number of successful AuthReqs; however, it does not provide an appropriate procedure for selecting the optimal CAC parameters. CAC-based authentication includes separate individual processing of contention-based exchanges of AuthReq, AuthResp, AssocReq, and AssocResp messages; it reportedly takes more than 60 seconds to register 2000 devices, and around 115 seconds for 3000 devices [5].

To minimize registration time, several approaches were proposed, such as the CAC method [5], [7]–[9], distributed authentication control (DAC) [10], and

a combined authentication/association (CAA) scheme [11]. Bankov *et al.* [8] proposed a contention-free transmission (CFT) scheme with several additional algorithms— Optimal Solution (OPT), Empty Slot Statistics (ESS), Decision Changing Algorithm (DCA), and Adaptive Threshold Algorithm (ATA)—to optimally adjust the authentication control threshold (ACT) value in order to reduce overall registration time. Bankov *et al.* [9] proposed a fast centralized authentication (FCA) scheme with up and down algorithms to select optimal CAC parameters. The usage of up/down algorithms significantly reduces authentication time for a large network with heavy background traffic.

The contention-based medium access control (MAC) protocols, however, generate severe collisions from massive registrations, and they provide poor performance at increased numbers of IoT devices. Ali *et al.* [12] proposed a smart grouping procedure to handle such limitations by grouping and splitting IoT devices into multiple domains. The smart grouping scheme was designed to identify devices that can be co-scheduled efficiently in densely deployed WLANs, but it mainly depends on the specific application. Moreover, the smart grouping mechanism is applicable only after the registration phase.

Liu *et al.* [13] designed a hybrid MAC for M2M networks that combines the benefits of both contention-based and contention-free protocols to address the performance issue in M2M communications. Their work improves the performance of data communications, compared to CSMA/CA and stand-alone TDMA, by introducing the *contention only period* (COP) and the *transmission only period* (TOP). In the COP, devices with a different contention probability contend for a transmission time slot, and thus, only successful devices are assigned a T-slot in the TOP. But device classification and prioritization to reduce contention are only available after registration.

In this paper, we propose an adaptive registration procedure for massive numbers of IoT devices in M2M networks. We propose a hybrid slotted-CSMA/CA-TDMA (HSCT) MAC protocol where a logical frame is divided into two parts: i) a contention-based slotted-CSMA/CA period (SCP) that is further divided into multiple CSMA/CA access windows (i.e., C-slots), and ii) a contention-free slotted-TDMA period (STP) that is further divided into multiple T-slots. The SCP is based on the CSMA/CA mechanism, which allows each device to select a backoff slot in a geometric probability distribution of the Sift slotted fixed-window CSMA protocol [14] to send its AuthReq. To ensure fairness, each device reselects a backoff slot within the SCP in each logical frame (i.e., each beacon interval). On the other hand, the STP is used to exchange AuthResp and AssocReq/AssocResp messages between devices and the AP through individually allocated T-slots without contents. To determine the optimal SCP/STP proportion, we propose an algorithm that maximizes the number of successful registrations. The proposed algorithm dynamically adjusts operational parameters to establish an efficient registration procedure in scenarios where large



FIGURE 1. Registration procedure in 802.11ah operation.

numbers of IoT devices simultaneously try to register at a single AP. The proposed HSCT scheme is based on the IEEE 802.11ah standard, which uses only a 1 \sim 2 MHz channel bandwidth, considering the low physical layer (PHY) transmission rate. The performance of the proposed HSCT was analyzed with Network Simulator 3 (NS-3) using a modified IEEE 802.11ah. Optimal HSCT (HSCTopt) consumes, on average, 64% and 87% less time compared to the existing optimal CFT (CFTopt) and CFT-ATA schemes. Simulation results demonstrate the effectiveness of the proposed HSCT MAC protocol.

The major contributions of this paper are as follows:

- A scalable hybrid MAC protocol incorporates efficient registration with dynamic slotted-CSMA/CA and slotted-TDMA mechanisms for large numbers of connected IoT devices.
- Two modified algorithms (smart-up and smart-down) select optimal CAC parameters.
- A CAC-based mechanism with a Sift geometric probability distribution reduces the number of contentions at the beginning of the SCP to mitigate collisions.
- Adaptive adjustment of the SCP and STP enables efficient channel utilization.
- A closed-form analytical model provides an average number of AuthReqs in the C-slots.

The remainder of this paper is organized as follows. Section II briefly introduces the background of the registration procedure under IEEE 802.11ah and some related work. The proposed HSCT MAC protocol is explained in Section III, and Section IV provides the performance evaluation of the proposed HSCT MAC protocol. Finally, Section V draws conclusions.

II. BACKGROUND AND RELATED WORK

A. REGISTRATION UNDER IEEE 802.11AH

Fig. 1 depicts the basic registration procedure under IEEE 802.11ah. We assume that registration includes both

AuthReq/AuthResp and AssocReq/AssocResp handshakes. The devices start the registration procedure after successful reception of a beacon, while the AP periodically broadcasts at the beginning of each beacon interval (BI). After reception of the beacon, all devices compete for the CSMA channel to send their AuthReqs to the AP. IEEE 802.11ah introduces two authentication mechanisms to mitigate severe contention during network initialization: centralized and distributed [3].

In the CAC method, the AP dynamically adjusts the ACT value (V_{ACT}) , which is included in every beacon frame to keep the number of requesting IoT devices at an optimum level. Wang [5] proposed a mechanism to adjust V_{ACT} whereby the AP adjusts V_{ACT} according to the length of the management queue (MQ) that buffers the response frames (i.e., AuthResp and AssocResp). A larger V_{ACT} allows more devices to send an AuthReq. Incrementing or decrementing V_{ACT} is decided upon after comparing the current MQ size (Q_L) with a fixed value for the queue size threshold (Q_T) . If Q_L is greater than Q_T , the AP considers the network congested and decreases V_{ACT} . On the other hand, if Q_L is less than Q_T , the AP considers the network under-loaded and increases V_{ACT} . However, it does not define the procedure to select optimal values for the MQ size, ACT value, and the increment/decrement step size, (Δ). The AP continues sending an updated V_{ACT} in subsequent beacons, and regulates the number of contending devices by adaptive adjustment of V_{ACT} . Devices receive an updated V_{ACT} in each BI and compare it with a uniform random number (U_R) , which is generated in the range [0, 1022] by each device during initialization. If $U_R \leq V_{ACT}$, then the device is allowed to send an AuthReq during the current BI. Otherwise, it is not allowed to access the channel until the next BI. According to the registration process, the association procedure starts after a successful AuthReq/AuthResp authentication procedure. The AP must properly control the number of devices that can successfully join during the current BI. For that reason, when the network is small, V_{ACT} should be higher to allow more devices to send an AuthReq to avoid unnecessary delays. On the other hand, in a large network, in order to limit the number of device requests, V_{ACT} should be lower to reduce the contention level. Consequently, the AP should dynamically select the optimal value for V_{ACT} based on the MQ size. After successful authentication, each device sends an AssocReq to the AP, which assigns an AID to the device and responds with AssocResp, which completes the registration procedure. Devices can exchange sensor/actuator data only after successful completion of the registration handshakes. This mechanism describes for choosing the optimal V_{ACT} , however, it provides the fixed Δ instead of the optimal selection that significantly affect the performance of the registration process as it is shown in [9].

In the DAC method [3], the AP periodically broadcasts beacons that include information about the network parameters: authentication control slot (ACS) duration (T_{ac}), minimum transmission interval (TI_{min}), and maximum transmission interval (TI_{max}). The default values of T_{ac} , T_{min} , and T_{max} are 10 time units (TUs), eight BIs, and 256 BIs, respectively. Each device maintains a transmission interval (*TI*) in BI units, and a *TI* is determined according to

$$TI_{r} = \begin{cases} TI_{min} & r = 0\\ min \{2TI_{r-1}, TI_{max}\} & 0 < r < R_{max} \end{cases}$$
(1)

where *r* is the number of authentication retries, and R_{max} is the maximum number of authentication retries. Each device chooses the number of BIs, *m*, from the uniformly distributed range $[0, TI_r - 1]$. Again, the BI is divided into *L* ACSs, where $L = T_{BI}/T_{ac}$. The ACS number, *l*, is uniformly selected from the range [0, L - 1]. The device attempts to send an AuthReq in ACS *l* of BI *m*. If the AuthReq attempt is unsuccessful, the device increases the number of authentication retries (*r*) and regenerates *m* and *l*. Inside the ACS, the device attempts to access the channel according to an enhanced distributed channel access mechanism.

Bankov *et al.* [9] studied the DAC method. Their work shows that the optimal selection of L minimizes the average registration time. The registration grows almost linearly with large numbers of connecting devices. Unlike CAC, DAC reduces the contention in a distributed way without knowing the traffic conditions. Since more in-depth study is necessary with mixed practices of registrations and data exchanges, error-channel conditions, and comparisons between CAC and DAC methods in the same environment, we leave that as future work. In this paper, we focus only on the CAC-based mechanism.

B. ENHANCEMENT OF THE REGISTRATION PROCEDURE

Sthapit *et al.* [7] proposed an analysis of CAC parameters. In that study, the total number of devices was divided into equal sub-groups, and only one of the sub-groups was selected for the association procedure. This scheme reduced the association time; however, the researchers did not provide any mechanism to prepare optimal sub-groups for any specific network size. Moreover, they assumed in the simulation that only small numbers of devices are active at one time, which is not a realistic scenario in a real M2M communications registration process for massive numbers IoT devices.

Bankov *et al.* [8] proposed a CFT scheme with several algorithms to adjust V_{ACT} in order to reduce overall registration time. Optimal Solution (OPT) provides the best performance among the CAC-based algorithms. An evaluation assumes the AP already has information on the total number of devices in the network. OPT, however, is less practical in the massive registration process in real IoT environments, because the AP cannot determine how many devices exist in the network until all devices are successfully registered. The adaptive threshold algorithm [8] provides significant improvement in device registration performance. It works in two phases (learning and working) to select the optimal V_{ACT} . ATA required a little bit more registration time than OPT [8].

One of our previous studies proposed an enhanced registration procedure with a network allocation vector (NAV) to mitigate contention in M2M communications [11], which reduces the time required for the registration process by implementing three policies: (i) use of a CAC method to limit the number of devices from the total number of devices in the M2M network, (ii) combined exchange of authentication and association messages to reduce contention, and (iii) extended AuthReq and AuthResp to carry NAV information in frames to allocate the communications channel to combine processing of authentication and association in order to minimize contention. In this registration procedure, only AuthReq faces contention to access the channel, and therefore, the combined scheme cuts contention by half, compared to the conventional 802.11ah. However, the performance of the combined scheme is not sufficient from a scalability perspective, and the registration process takes rather a long time. Even though the scheme provides an updated NAV value in the AuthResp frame, the network becomes congested as it grows.

All the above-mentioned studies have several limitations, including regeneration of random values in each BI and a fixed step size (Δ). The current version of IEEE 802.11ah calls for the random value to be generated at initialization, and it can be regenerated only after a successful authentication procedure. The devices generate their random value only once, and therefore, the AP can eventually increase V_{ACT} up to its maximal value to ensure that all the devices have started authentication. The use of a fixed step size is inefficient, depending on different sizes of networks [9].

Bankov *et al.* [9] proposed the FCA scheme's up and down algorithms, which adaptively adjust the authentication control threshold and step size based on the management queue size. They compared the performance of the up and down algorithms with the optimal condition under the assumption that the exact number of devices to be authenticated is known. In a simulation with realistic conditions, the up and down algorithms provided authentication times just exceeding 22% and 20% more than the optimal condition, respectively.

All the above approaches are contention-based CSMA/CA mechanisms, and due to the proportional increase in contention with increased number of devices, there is a limitation in performance enhancement for CSMA-based large-scale M2M networks. In this paper, we propose a dynamic HSCT MAC protocol that supports efficient exchange of management frames for large-scale M2M networks with massive numbers of IoT devices by utilizing the benefits of both contention-based CSMA and reservation-based TDMA. Both smart-up and smart-down algorithms are useful not only for HSCT but also for CFT or for any CSMA/CA-based scheme for the optimal selection of CAC parameters. Moreover, we provide a comprehensive analysis of the proposed HSCT MAC scheme, and compare it with the traditional IEEE 802.11ah access schemes in an M2M network.

III. HYBRID SLOTTED-CSMA/CA-TDMA SCHEME

A. PRELIMINARY ASSUMPTION

We consider an M2M network with N IoT devices and an AP as the network coordinator in a star topology, where



FIGURE 2. Registration procedure in the hybrid slotted-CSMA/CA-TDMA scheme.

each device can transmit and receive management frames from the AP. The star topology can be found in various IoT applications, including industrial and agriculture monitoring, home/building automation, healthcare system monitoring, and smart metering [15]. Each device is identified by its MAC address and a unique AID that is assigned after the registration process. It is assumed that each device always intends to connect with the AP. Fig. 2 depicts a detailed sequence diagram inside the beacon interval.

In the network, transmission time is divided into a constant period BI, denoted by T_{BI} , which is composed of three parts: beacon period (BP), slotted-CSMA/CA period, and slotted-TDMA period, where the durations are denoted by T_{BP} , T_{SCP} , and T_{STP} , respectively, as shown in Fig. 3. The AP broadcasts beacon frames to all devices in each BP to inform the devices of the following: the beginning of the SCP; the number of total C-slots (*num_C_Slots*) in the SCP; the duration of each C-slot; the number of total T-slots (*B*), along with their durations; the authentication control threshold value; the MAC addresses for non-registered devices; and AIDs for registered devices that are allowed to use the T-slots.

B. ACCESSING THE C-SLOT WINDOW

In the HSCT MAC, devices transmit an AuthReq within the SCP of each BI according to the dynamic slotted-CSMA/CA contention protocol, in which devices attempt to send the



FIGURE 3. The frame structure in the hybrid slotted-CSMA/CA-TDMA scheme.

AuthReq according to the CSMA/CA backoff process with a fixed number of backoff slots of duration δ [14]. The operation of dynamic slotted-CSMA/CA is different from the traditional CSMA/CA protocol. First, the contention period is divided into multiple mini-CSMA/CA slots (C-slots); therefore, only one group of specific devices is allowed access using their specific C-slot windows. Second, the AuthReq transmissions are permitted only during the SCP; thus, the devices that are not allowed access in the current BI must wait for the next SCP in the next BI. Third, a backoff slot is randomly selected from a fixed range of contention windows (K) of duration δ with a Sift distribution within the C-slot. Finally, before each device attempts an AuthReq transmission when the backoff counter becomes zero, it is required to check whether the remaining time in the current access window is long enough to complete the AuthReq message exchange, including authentication request acknowledgment (AuthReqAck), and the related distributed inter frame-space (DIFS), short inter-frame space (SIFS), and guard time (T_{GT}) . Fig. 3 shows the structure of the slotted-CSMA/CA access mechanism.

The IEEE 802.11 standard defines two link-level authentication types: *open system* and *shared-key* [21]. In the open system, the authentication procedure consists of two frame exchanges, and the AP accepts the device after verification of identity. The AuthResp frame has some fixed-length fields, and four frame exchanges are required, including the association request/response. On the other hand, sharedkey authentication consists of four management frames of subtype authentication, and the AuthResp frame has some variable-length fields. Therefore, six frame exchanges are required to complete the registration process. Considering these different link-level authentication types, the AP should be able to provide different T-slot sizes according to the type of authentication algorithm. In this paper, we only consider the open system authentication.

During the registration stage, all devices belong to either (i) the *access group*, comprising devices that are allowed to send an AuthReq in the current SCP, or (ii) the *deferred group*, comprising devices that must wait for the next SCP. The registrations of the access group are handled using the CAC method. Devices randomly select one C-slot from all CSMA/CA slots defined in the beacon frame. It is expected that higher numbers of successful registrations can be obtained if (i) contention is at an optimum level with an optimal number of registration requests, and (ii) the STP is long enough to handle the optimal number of registration requests. By providing more time for the SCP in the fixed BI, the duration of the STP will be decreased, and the transmission time for successful devices is reduced. Thus, there is a tradeoff between the durations of the SCP and the STP in a fixed BI.

From each beacon, the devices can extract V_{ACT} , num_C_Slots, and their durations. Massive contention can be mitigated by Algorithms 1, 2, and 3. First, the CAC method reduces the number of devices by configuring the access group that allows accessing the channel using contention-based C-slots. Secondly, the access group is partitioned into the randomly selected C-slots (C - Slots = $1, \ldots, num_C_Slots$), where only one C-slot is uniformly selected from *num_C_Slots* in a BI. The devices are configured according to the C-slot information (i.e., C-slot number, and C-slot duration) in order to send an AuthReg. The backoff procedure works only in the assigned C-slot determined by the slot number, as well as the C-slot with the starting time and ending time in the backoff mechanism. Therefore, only an optimal number of devices contend for the channel in their own C-slot, and they stay in energy-saving mode during other C-slots in the SCP. If multiple devices try to send AuthReq frames simultaneously, collisions may occur, and the backoff counter is reset to resolve collisions. Once a successful C-slot is obtained, the AuthReq frame is transmitted to the AP.

C. ADAPTIVE CONTROL OF CAC PARAMETERS

The up and down algorithms in FCA [9] provide efficient performance in the authentication procedure. Both algorithms select V_{ACT} and Δ adaptively, based on the management queue size and with the execution of three modes (waiting, studying, and working). In the authentication procedure, Q_L is obtained from the buffers of the response frames (i.e., AuthResp) at the AP. However, in the registration procedure, Q_L is obtained from the buffers of both response frames (i.e., AuthResp and AssocResp). Since the traffic is composed of not only the response frames but also the request frames (i.e., AuthReq and AssocReq), the optimal selection of V_{ACT} and Δ should therefore consider the overall traffic. We developed smart-up and smart-down algorithms, as shown in Algorithm 1 and Algorithm 2, which extend the up and down algorithms [8]. Both smart-up and smart-down algorithms have the same three modes (*waiting*, *studying*, and working) as the up and down algorithms; however, the working principal is different. Both smart-up and smartdown algorithms have the same waiting mode, and the AP initially executes in waiting mode where if Q_L is empty,

VOLUME 6, 2018

Algorithm 1 Smart-Up Algorithm for Optimal Value Selection of V_{ACT} at AP

// VACT: the authentication threshold value // Δ : step size of ACT *II init stage*: flag that is used to find more optimal Δ // maxACT: maximum authentication value // Q_L : management queue size of the AP // SA: number of successful AuthReq/AssocReq in the previous BI // mode: assign the specific mode *II change*_ Δ : flag that is used to set more precise Δ 1: set maxACT = 1023, $V_{ACT} = 0$, $init_stage = 1$ 2: set $\Delta = 0$, change $\Delta = 0$, mode = Waiting Mode while BeaconInterval(BI) do 3: 4: $Q_L = GetMgtOS()$ $S_A = GetNumAuthReqInBI()$ 5: **if** mode == WaitingMode **then** 6: 7: Do waiting mode () // waiting mode 8: else if *mode* == *StudyingMode* then 9: Do studying mode () // studying mode 10: else if mode == WorkingMode then 11: *Do_working_mode ()* // working mode 12: end if 13: end while procedure Do_waiting_mode () 1: 2: if $Q_L == 0$ then 3: if *init_stage* == 1 then 4: $V_{ACT} = 0.5 \times maxACT$ 5: $init_stage = 0$ 6: else 7: $V_{ACT} = maxACT$ 8: end if 9: else 10: $mode = StudyingMode, V_{ACT} = 1, \Delta = 1$ 11: end if 12: end procedure procedure Do_studying_mode () 1: 2: if $Q_L == 0$ then 3: if $S_A == 0$ then $\Delta = 2\Delta$ 4: 5: else 6: $\Delta = \Delta + 1$ 7: end if 8: $V_{ACT} = V_{ACT} + \Delta$ 9: if $V_{ACT} \ge maxACT$ then 10: Do waiting mode () 11: end if 12: else 13: if $\Delta > 1$ then 14: $\Delta = \Delta/2$ 15: end if 16: $change_{\Delta} = 1, mode = WorkingMode$ 17: end if

18: end procedure

Algorithm 1 (Continued.) Smart-Up Algorithm for Opt	imal
Value Selection of V_{ACT} at AP	

1:)	procedure Do_working_mode ()
2:	if $Q_L == 0$ then
3:	if $change_{\Delta} == 1$ AND $S_A == 0$ then
4:	$\Delta = \Delta + 2$
5:	else if $(change_{\Delta} == 1$ AND $S_A! = 0)$ OR
6:	$(change_{\Delta} == 0$ AND $S_A == 0$) then
7:	$\Delta = \Delta + 1$
8:	end if
9:	$V_{ACT} = V_{ACT} + \Delta$
10:	if $V_{ACT} \ge maxACT$ then
11:	Do_waiting_mode ()
12:	end if
13:	else
14:	$change_{\Delta} = 0$
15:	end if
16:	end procedure

the AP sets V_{ACT} to the half of its maximum value (i.e., $0.5 \times maxACT$) instead of the maximum value used in both up and down algorithms. This value is assigned only during initialization of waiting mode, it reduces the convergence time to switch from waiting mode to studying mode. When Q_L becomes nonempty, the AP changes its state to studying mode and initializes the parameters.

For smart-up in studying mode, the AP increases V_{ACT} and Δ based on both management queue size and the number of successful AuthReq/AssocReq handshakes in the previous BI (S_A) to find an optimal Δ in order to register in each BI as many devices as possible. The smart-up algorithm initializes V_{ACT} and Δ to 1. In studying mode, if the management queue is empty $(Q_L == 0)$ and there was no AuthReq/AssocReq handshake in the previous BI ($S_A == 0$), then the AP considers the current Δ to be too low, and therefore, Δ is doubled. On the other hand, if Q_L is empty, but S_A is nonzero, then the AP increases Δ by 1, as in normal mode. Although the queue is empty, however, devices can send AuthReq/AssocReq messages using the current Δ . In this case, if Q_L is nonempty, the AP switches to working mode. In working mode, Δ is only updated if Q_L is empty. Moreover, if flag change_ $\Delta == 1$ and S_A is zero, then Δ is increased by 2; otherwise, it is increasesd by 1. The differences between the smart-up algorithm and the up algorithm are depicted in lines 3-8 of *Do_waiting_mode()*, lines 4-8 of Do_studying_mode(), and lines 4-9 of Do_working_mode() in Algorithm 1.

In the smart-down algorithm, the AP sets V_{ACT} to the half of maximum value (i.e., $0.5 \times maxACT$) if Q_L is empty; otherwise, V_{ACT_old} is set to V_{ACT} , and V_{ACT} is set to 0 upon initialization of studying mode. In this mode, if Q_L is nonempty, or if S_A is nonzero, V_{ACT} is set to $0.5 \times V_{ACT_old}$, and V_{ACT_old} is set to V_{ACT} to reduce traffic. On the other hand, if Q_L is empty, and if S_A is zero, Δ is set to V_{ACT} ,

and V_{ACT} is updated as $V_{ACT} + \Delta$. Then, the AP switches to working mode, and the parameters are updated according to the same $Do_working_mode()$ procedure in Algorithm 1. The differences between the smart-down and down algorithms are depicted in lines 3-8 of $Do_waiting_mode()$ and in lines 4-8 of $Do_studying_mode()$ in Algorithm 2.

Algorithm	3	Scheduling	Authentication	Request	With
Authenticati	on	Control Thre	eshold (ACT) Va	lue at Dev	ice

$//V_{min}$: Minimum value of the range of the random value
$//V_{max}$: Maximum value of the range of the random value
// C _{Slot} : CSMA/CA slot number
$//U_{P}$ a random value by the uniform distribution

 $1/U_R$: a random value by the uniform distribution

1:	set $U_R = (int) random [V_{min}, V_{max}]$
	$//V_{min} = 0, V_{max} = 1022$
2:	loop
3:	wait until Beaconisreceived
4:	set $C_{Slot} = (int) random [0, GetTotalCsmaSlot()]$
5:	set $V_{ACT} = GetACT()$
6:	if $U_R \leq V_{ACT}$ then
7:	schedule AuthReq at C_{Slot} access slot
	// access group
8:	else
9:	wait for the next Beacon // deferred group
10:	end if
11:	end loop

The device uses Algorithm 3 to compare V_{ACT} with random value U_R , which is selected from a *uniform* distribution. If $U_R \leq V_{ACT}$, the device belongs to the access group; otherwise, it belongs to the deferred group.

D. SIFT GEOMETRIC PROBABILITY DISTRIBUTION FOR BACKOFF SLOT SELECTION

Devices in IEEE 802.11ah use CSMA/CA, where the backoff slot is randomly chosen in the contention resolution procedure based on a uniform distribution that provides the same probability that a device will collide in any one slot. Although every device has an equal opportunity to pick one of the K backoff slots, the network experiences high contention at the beginning of each C-slot window. Tay *et al.* [14] proposed the slotted fixed-window CSMA protocol known as Sift for wireless sensor networks (WSNs). The Sift probability distribution function (p_r) is defined as the probability of selecting the *r*-th backoff slot, which is determined as follows:

$$p_r = \frac{(1-\alpha)\,\alpha^K}{1-\alpha^K} \cdot \alpha^r, \quad r = 1, \dots, K$$
(2)

where $\alpha = M^{-1/(K-1)}$ is a distribution parameter, $0 < \alpha < 1$, and M is the maximum number of contenders. Fig. 4 compares the selection probability for backoff slot r in uniform and Sift geometric probability distributions. In the Sift distribution, the probability of selecting backoff slots from the front is lower than the selection of backoff slots from the end. The advantage is that if large numbers of devices in the network simultaneously try to access a slot, only a few devices select the front backoff slots and the others select latterly positioned slots. Therefore, the probability of collision in the front backoff slots is reduced, and more successful transmissions are possible. As with the initialization of the



FIGURE 4. Probability of selecting backoff slot (r) based on uniform and *Sift* distribution for K = 32 and different values for M.



$$\begin{split} T_{C}^{i} &= T_{AuthReq} + T_{AuthReqAck} + 2SIFS + T_{GT} \\ T_{AuthTransmit} &= DIFS + T_{AuthReq} + SIFS + T_{AuthReqAck} \\ t_{e}^{i} &= t_{s}^{i} + T_{i}, t_{1}^{i} = t_{e}^{i} - T_{AuthTransmit} - T_{C}^{i} \\ t_{2}^{i} &= t_{e}^{i} - T_{AuthTransmit} - T_{GT}, t_{3}^{i} = t_{e}^{i} - T_{C}^{i}, t_{4}^{i} = t_{e}^{i} - T_{V}^{i} \\ t_{5}^{i} &= t_{e}^{i} - SIFS - T_{GT} \end{split}$$



contention window, the devices use 32 as the value of both K and M in our experiment.

E. ANALYSIS OF THE ACCESS PERIOD IN A CSMA-SLOT

Here, we formulate a closed-form analytical model of the expected number of AuthReqs ($E[\Gamma_A^i]$) in the *i*-th C-slot, as depicted in Fig. 5. The analysis of HSCT is inspired by the analytical model proposed by Zhang *et al.* [16] for a reservation- and contention-based hybrid MAC for WLANs. In the analysis, the channel is assumed to be error-free, and collision occurs when more than one device simultaneously transmits frames. The approach is based upon the equilibrium conditions of the IoT network for each single IoT device, where the expected values of the system variables must satisfy the given relationship among them, and all IoT devices are statistically equivalent due to the fairness nature of the CSMA/CA-based MAC protocol. These assumptions result in the same statistics for all IoT devices.

As depicted in Fig. 5, the *i*-th C-slot period, T_i , starts from t_s^i and ends at t_e^i , and the devices can perform backoff within the *backoff period* interval $T_B^i = T_i - T_{GT}$. The backoff period (T_B^i) $[t_s^i, t_e^i - T_{GT}]$ in the *i*-th C-slot is further divided into the access period (T_A^i) and the vulnerable *period* (T_V^i) . If a device is selected as the access group, the transmitter needs to make sure that the whole duration of AuthReq message exchanges, including AuthReqAck, can be completed before $t_5^i = t_e^i - SIFS - T_{GT}$; otherwise, the frame transmission may incur a conflict with the next C-slot or T-slot, according to the ECMA-368 standard [17]. Therefore, after an acceptable transmission attempt, the remaining time should be greater than the conflict period $[T_C^i]$ = $T_{AuthReq} + T_{AuthReqAck} + 2SIFS + T_{GT}$], where $T_{AuthReq}$ is the time to transmit one AuthReq frame, $T_{AuthReqAck}$ is the time to receive the acknowledgment (ACK) for a transmitted AuthReq, the SIFS is the short inter-frame space before receiving the ACK, and T_{GT} is the guard time. If the backoff counter reaches zero within the conflict period, the conflict can be resolved by either (i) the *hold-on strategy*, where the device just holds on and transmits during the next SCP, or (ii) the *backoff strategy*, where the device invokes another stage of the backoff procedure [16], [18]. It has been shown that the backoff strategy provides better performance with a moderate traffic load [19], and therefore, it is adopted in this paper.

In the backoff strategy, the backoff procedure is performed in time interval T_{R}^{i} (the backoff period); however, transmission is allowed only up to the access period interval T_A^i . If an AuthReq transmission is initiated within the interval $[t_s^i, t_2^i = t_e^i - T_{AuthTransmit} - T_{GT}]$, where $T_{AuthTransmit} =$ $DIFS + T_{AuthReq} + SIFS + T_{AuthReqAck}$, the AuthReq is transmited either successfully or with a collision, according to the CSMA/CA mechanism. The packet transmission is executed within the interval $[t_s^i, t_e^i]$, and there can be an ineffective interval, T_V^i (the vulnerable period), between the last AuthReq transmission finishing point (i.e., t_4^i) and the end of the *i*-th C-slot, T_i (i.e., t_e^i). If an AuthReq transmission attempt is suspended in the current C-slot of the SCP, then a new AuthReq is initiated, and the backoff slot is reset in the next SCP after making a virtual grouping. Other devices that are not involved in the conflict period pause their backoff counters at the starting point of the backoff period (T_R^i) .

Let E[B] denote the expected number of backoff slots, and let E[C] denote the transmission trials experienced by one AuthReq frame. When a device is busy (in backoff stage or in transmission), the transmission probability (τ) for each device in any time slot is expressed as follows [16]:

$$\tau = \frac{E[C]}{E[B] + E[C]}$$
(3)

In Eq. (3), transmission probability τ can be approximated, based on renewal reward theory, as a ratio of the average reward received during a renewal cycle over the averAlgorithm 4 Backoff Slot Selection Algorithm of Every Device n_i $//n_i$: the AID (Association Identifier) of the *i*-th device $//\alpha$: a distribution parameter $//l_i$: a random number generated by the *i*-th device *IIM*: the maximum number of contenders //K: the size of the contention window $//r_i$: the Backoff slot number at the *i*-th device *Ilisten*: the sensing of channel by the *i*-th device 1: while (1) 2: wait until new Beacon frame 3: extracts the K and M value from Beacon frame 4: Retry: set $r_i = Geometric$ distribution value (K, M) 5: **set** $listen_i = IsChannelBusy()$ 6: 7: if ChannelIsBusy() then 8: wait for the channel free longer than DIFS 9: end if 10: while $r_i > 0$ do 11: if $r_i == 0$ then 12: n_i sends control frame to the AP 13: if CollisionOccured() then 14: Go to Retry 15: end if 16: end if 17: if ChannelIsBusy() then 18: freeze r_i and wait for channel free longer than DIFS 19: end if 20: $r_i = r_i - 1$ 21: end while 22: end while 1: procedure Geometric_distribution_value (K, M) set $\alpha = M^{-(1)/(K-1)}$ 2: 3: while (1) do 4: set $l_i = random [0, 1]$ 5: for r = 1 to K do if $p_r = ((1 - \alpha)\alpha^K / 1 - \alpha^K)\alpha^{-r} > l_i$ then 6: 7: return r 8: end if 9: r = r + 110: end for 11: end while 12: end procedure

age length of the renewal cycle [16], [20]. If a transmitted AuthReq frame collides with probability P_{fc} , the expected number of transmission trials follows a truncated geometric distribution with success probability $(1-P_{fc})$. Then, E[B] and E[C] are expressed as

$$E[B] = \sum_{\substack{s=1\\s}}^{S} (E[b_s]) \cdot P_{fc}^{s-1}$$
(4)

$$E[C] = \sum_{s=1}^{S} P_{fc}^{s-1}$$
(5)

where $E[b_s]$ is the expected number of backoff slots in the s-th backoff stage, and S is the maximum number of backoff stages. Since HSCT follows a geometric distribution–based probability distribution, as defined in Algorithm 4, $E[b_s]$ in Eq. (4) is defined as

$$E[b_s] = \sum_{r=1}^{K} r \cdot p_r \cdot (1 - p_r)^{r-1}$$
(6)

where p_r is the Sift probability for selection of backoff slot r. Then, p_r can be determined using Eq. (2). Hence, Eq. (3) can be rewritten as

$$\tau = \frac{\sum_{s=1}^{S} P^{s-1}}{\sum_{s=1}^{S} (E[b_s]) \cdot P^{s-1} + \sum_{s=1}^{S} P^{s-1}}$$
(7)

The backoff (mini) slots in T_A^i may have the following three states.

a) The slot may be *idle* if there is no AuthReq transmission by any device, with backoff slot duration δ , and the probability can be expressed as

$$a_A^i = (1 - \tau)^{N_{ac}}$$
(8)

b) The last time instant when the AuthReq transmission attempt can be allowed is $t_3^i = (t_e^i - T_C^i)$, and conflict period T_C^i is smaller than the duration of a successful AuthReq transmission frame $(T_{AuthTransmit})$. Since the whole transmission for AuthReq messages $(T_{AuthTransmit})$ should finish before t_5^i , that is, at least one SIFS plus one guard time (T_{GT}) before the end of the *i*-th C-slot, if an AuthReq transmission starts in the interval $[t_2^i, t_3^i]$, the AuthReq transmission will continue until the next C-slot/T-slot. Thus, the duration of such a transmission $(T'_{AuthTransmit})$ is less than $T_{AuthTransmit}$. For such a duration, the beginning of the frame transaction is uniformly distributed inside $[t_2^i, t_3^i]$, with an expected duration of

$$E\left[T_{AuthTransmit}^{'}\right] = \frac{T_{AuthTransmit} + T_{C}^{i}}{2}$$
(9)

Since the AuthReq should start before the conflict period (i.e., $T_B^i - T_C^i + T_{GT}$), the $T_{AuthTransmit}^{'}$ can be defined as $\frac{T_{AuthTransmit} - T_C^i}{T_B^i - T_C^i + T_{GT}}$, and the probability of the restricted slot (slots in the interval $[t_2^i, t_3^i]$) can be expressed as

$$b_A^i = \left(1 - a_A^i\right) \cdot \frac{T_{AuthTransmit} - T_C^i}{T_B^i - T_C^i + T_{GT}}$$
(10)

 c) The backoff (mini) slots may be contained in a AuthReq transmission as either *successful* or *collided*. Therefore, the probability of a successful/collided slot is expressed as

$$c_A^i = 1 - a_A^i - b_A^i \tag{11}$$

Let D_A^i be the duration of a generic slot inside T_A^i ; then, the expected duration of a generic slot within access period

$$E\left[D_{A}^{i}\right] = a_{A}^{i} \cdot \delta + b_{A}^{i} \cdot T_{AuthTransmit}^{'} + c_{A}^{i} \cdot T_{AuthTransmit}$$
(12)

On the other hand, if the last AuthReq transmission is initiated within time interval $[t_1^i, t_3^i]$, such a transmission ends at point $t_4^i(t_e^i - T_V^i)$, which is positioned in conflict period T_C^i , where vulnerable period T_V^i (denoting the idle duration) is shorter than the conflict period T_C^i , on average. In this case, we estimate that the starting point of the transmission is uniformly distributed inside $[t_1^i, t_3^i]$, with expected vulnerable period $T_V^i = \frac{T_C^i}{2}$. However, if there is no transmission within this interval, then $T_V^i = T_C^i$. The number of idle backoff slots in $[t_1^i, t_3^i]$ can be defined as

$$\Gamma_{T_{AuthTransmit}} = \frac{T_{AuthTransmit}}{\delta}$$
(13)

and the probability of no transmission in this interval is $a_A^{i\Gamma_{T_{AuthTransmit}}}$. Thus, the expected length of T_V^i can be expressed as

$$E\left[T_{V}^{i}\right] = a_{A}^{i\Gamma_{AuthTransmit}} \cdot T_{C}^{i} + \left(1 - a_{A}^{i\Gamma_{AuthTransmit}}\right) \cdot \frac{T_{C}^{i}}{2}$$
$$= \left(1 + a_{A}^{i\Gamma_{AuthTransmit}}\right) \cdot \frac{T_{C}^{i}}{2}$$
(14)

Now, the expected length of the access period within a C-slot can be obtained with $E[T_A^i] = T_i - E[T_V^i]$, as shown in Fig. 5, and the expected number of AuthReqs in a C-slot can be derived from

$$E\left[\Gamma_{A}^{i}\right] = \frac{E[T_{A}^{i}]}{E\left[D_{A}^{i}\right]}$$
(15)

Solving Eq. (15) to find out the number of expected transmission slots within a T_A , we require a numerical solution for two unknowns: τ and P. Collision probability P obtained as follows since the collision can only occur if any other station also transmits in the same time slot outside the vulnerable (conflict) period. Since the probability of transmitting for each node is τ , the collision probability can be written as

$$P = 1 - (1 - h)(1 - \tau)^{N_{ac} - 1}$$
(16)

where h is the probability for a generic slot to be in the vulnerable time, and that can be obtained as

$$h = \frac{\Gamma_{\rm V}}{\Gamma_{\rm A} + \Gamma_{\rm V}} \tag{17}$$

where the average number of slots, Γ_V , within a vulnerable period can be estimated as $\Gamma_V = \frac{T_V}{\delta}$, with Γ_A already defined in Eq. (15). Therefore, the expected number of AuthReqs in one SCP can be derived as

$$E\left[\Gamma_{\text{SCP}}\right] = \left\lfloor E\left[\Gamma_A^i\right]\right\rfloor \cdot num_C_Slots \tag{18}$$

where *num_C_Slots* is the total number of C-slots, and the duration of all the C-slots is the same.



FIGURE 6. Average number of AuthReqs in one SCP (A_{avg}) and actual contenders (N_{ac}) with different step size (Δ)(num_C_Slots = 3).



FIGURE 7. The format of the registration information (RI) field.

The average number of successful AuthReqs in one SCP (A_{avg}) and the actual number of contenders (N_{ac}) with a different step size (Δ) for a Sift geometric probability distribution is plotted in Fig. 6. The step size provides the actual number of contenders (N_{ac}) among the N devices, defined as

$$N_{ac} = \frac{\Delta}{N} 1023 \tag{19}$$

It is possible to get the average number of successful AuthReqs in one SCP (A_{avg}) providing a different number for N_{ac} , selected with Δ , as in Eq. (19). We can find the maximum number of successful AuthReqs in one SCP (A_{opt}) by providing an optimal step size, $\Delta_{opt} = 3$ for 8000 devices.

We can see that both the analytical and the simulation results give similar performance for all numbers of contenders.

F. ACCESSING TDMA SLOTS

The AP replies with AuthReqAck after successful reception of an AuthReq. The AuthReqAck is different from the traditional ACK, in that T-slot information (i.e., the T-slot index of the device) is included. Therefore, the devices forward the AuthReqAck frame from the MAC low layer to the MAC high layer to extract the T-slot information at the devices. Moreover, the AuthReqAck message prevents the devices from retransmitting the AuthReq in the current SCP. Before providing T-slot information, the AP checks whether T-slots are available or not. If a T-slot is available, the T-slot index is included; otherwise, the AP sends a "no free T-slot" status flag that acknowledges successful reception of the AuthReq with no available T-slot from the AP. The status flag tells the device to wait for the STP in the next BI. On the other hand, if the total number of AuthReqs in the current and previous logical frames is more than *B*, the AP queues the AuthReqs in the buffers and addresses them in first-in-first-out (FIFO) order in the subsequent BI. Upon receiving AuthRegAck from the AP, the device will stop sending the AuthReq and will wait for the allocated T-slot. However, if the device does not receive an AuthReqAck within the AuthReqAck timeout limit, the device resends the AuthReq through the current C-slot. The beacon frame contains a registration information (RI) block that consists of several RI fields, as shown in Fig. 7. If the device verifies its MAC address in the RI field (RIF), the device will get a TDMA slot number (TSN) in which to exchange the remaining frames in the assigned T-slot. The total number of RIFs is less than or equal to B T-slots.

The STP is divided into B T-slots with fixed duration, where each T-slot allows the exchange of AuthResp/ AssocReg/AssocResp with ACK frames. Thus, the AP replies with AuthResp following a SIFS. The successful device turns on the radio channel in its assigned T-slot to receive the AuthResp and sends the AssocReq to the AP using its allocated T-Slot; it can turn its radio channel off at all other times to save energy. Through the specified T-slot, the AP receives the AssocReq without collision and replies with an AssocResp frame that provides the capabilities information, AID, and supported rates, which is necessary information for the device. In case of hidden stations and high noise in the channel, however, it is possible for the AP or a device to fail to deliver one of the frames while participating in registration during the T-slot. Usually, standards designate AuthenticationRequestTimeout and AssociationRequestTimeout for resending an AuthReq and an AssocReq, respectively, after the request timeout has occurred. Therefore, if devices cannot successfully receive an AuthResp within the timeout limit, the device resends the AuthReq. On the other hand, if a device cannot successfully receive an AssocResp, the device sends a power save poll (PS-poll) request to get the AP to assign a T-slot.

G. ADAPTIVE CONTROL OF SCP AND STP DURATION

For efficient resource utilization, durations for the SCP and the STP must be adaptively adjusted according to the traffic load of the IoT network. The AP should provide a sufficient length of time for the SCP that ensures more successful transmissions of AuthReqs, as well as an optimal duration for

IEEEAccess

Algorithm	5	Procedure	for	Adjusting	Slotte	ed-CSMA/CA
Period (SCI	2) ;	and Slotted-	-TDI	MA Period	(STP)) Duration

 $//D_{SCP \ curr}$: SCP duration of the current cycle //D_{SCP prev}: SCP duration of the previous cycle $// D_{BF}$: the duration of the Beacon frame $//T_{BI}$: Beacon Interval (BI) $//D_{STP \ curr}$: STP interval of the current cycle //Rtotal: total number of slot requests (data or Management frames) $//R_{a prev}$: number of AuthReqs in previous BI $//R_{p \ prev}$: number of PsPollReqs in previous BI $//B_{mgt}$: number of T-slots for management (mgt.) frame in a STP $//B_{data}$: number of T-slots for data frame in a STP $//\theta$: unit of the increment (in milliseconds) 1: while (1) 2: set $R_{total} = GetTotalSlotReqInOneBI()$

3: set $B_{mgt} = GetTotalMgtTdmaSlotInOneBI()$ 4: set $B_{data} = GetTotalDataTdmaSlotInOneBI()$ 5: set $R_{rem} = GetNumberOfExtraMgtReq$ () 6: if $R_{total} == 0$ then 7: $D_{SCP \ curr} = 0.5 \times T_{BI}$ 8: else 9: if $(R_{a_prev} + R_{p_prev}) > (B_{mgt} + B_{data})$ then 10: $D_{SCP \ curr} = D_{SCP \ prev} - \theta$ 11: else if $(R_{a_prev} + R_{p_prev}) < (B_{mgt} + B_{data})$ then 12: 13: $D_{SCP \ curr} = D_{SCP \ prev} + \theta$ 14: else 15: $D_{SCP_curr} = D_{SCP_prev}$ 16: end if 17: end if 18: end if $D_{STP \ curr} = T_{BI} - D_{SCP \ curr} - D_{BF}$ 19: 20: end while

the STP that successfully executes the remaining handshakes in the registration process. Therefore, if the length of the SCP is inadequate (i.e., it is unable to allow an appropriate number of AuthReqs that is less than the number of T-slots in the STP), then some T-slots may be unused, and there may be some waste. On the other hand, increasing the duration of the SCP allows the number of AuthReqs to exceed the number of T-slots in the STP; then, after allocating all T-slots in the current SCP, the remaining devices must wait for the next BI. This reduces the number of T-slots, and therefore, channel utilization is decreased in the registration process.

Algorithm 5 depicts the procedure to adjust for optimal durations of the SCP and STP. The AP tries to maintain an equal number of requests and B ($B = B_{mgt} + B_{data}$) T-slots. However, it is possible that there are more requests than T-slots in a BI. In this case, the AuthReqs/PS-poll requests in the previous logical frame always get higher

TABLE 1. MAC layer parameters used in simulations [3], [24].

Parameters	Value
AP's transmission range	1000 meters
Device's transmission range	500 meters
Device position	random
PHY header + preamble	20 µs
Beacon length	67~100 bytes
Authentication request length	26 bytes
Authentication response length	24 bytes
Association request length	37 bytes
Association response length	27 bytes
Authentication request timeout	500 ms
Association request timeout	500 ms
AuthReqAck/ ACK length	14 bytes
Backoff slot duration (δ)	52 μs
SIFS	160 μs
DIFS	264 μs
Air propagation delay	3 μs
Guard time (GT) and IFG	50 µs
Beacon interval	100 ms
Backoff slots (<i>K</i>)	32
Retry limit	7

priority than the AuthReqs/PS-poll requests in the current logical frame. To enhance channel utilization of the SCP and STP, the duration of the SCP should be optimally adjusted to achieve an ideal balance, i.e., the number of successful AuthReqs in the SCP should be equal to the number of T-slots in the STP within a BI. The AP calculates the optimal value for the SCP duration based on collection of the current traffic in AuthReqs/PS-poll requests.

IV. PERFORMANCE EVALUATION

A. SIMULATION CONFIGURATION PARAMETERS

We modified the implementation of IEEE 802.11ah MAC and the PHY in the NS-3 [22] prepared for 802.11ah in version 3.24 [23]. The analysis and performance comparisons of IEEE 802.11ah [5], optimal CFT (CFTopt) and contention-free transmission-adaptive threshold algorithm (CFT-ATA) [8], and the CAA procedure [11] combined with the proposed HSCT procedure were provided under identical operational conditions. Considering the low-power nature of battery-powered sensors in IoT applications, transmission power was limited to 3 dBm. The payload size of the packets was 128 bytes for interfering devices. A simulation was performed with a 650 Kbps physical data rate using a 2 MHz channel bandwidth, and the PHY and MAC layer parameters were configured according to the IEEE 802.11ah draft [3], [24] and the IEEE 802.11 standard [21] as listed in Table I and Table II. We analyzed the optimal configurations of the ACT-based contention parameters with different num_C_Slots, interference traffic, and distributions. The AP handled up to 8000 devices randomly placed in a circle around it within a distance of, at most, half of the transmission range of the AP to avoid the influence of the hidden node problem [25], [26]. All the simulation results were averaged over 10 runs.

TABLE 2. Physical layer parameters used in simulations [3], [24].

Parameters	Value
Carrier frequency	900 MHz
Channel bandwidth	2 MHz
Physical rate	650 Kbps
Transmission power (uplink)	30 dBm
Transmission power (downlink)	30 dBm
Transmission gain (uplink)	3 dBi
Transmission gain (downlink)	3 dBi
Reception gain (uplink)	3 dBi
Reception gain (downlink)	3 dBi
Noise figure (uplink)	3 dB
Noise figure (downlink)	5 dB
Propagation loss model	Outdoor, macro [23]
Error rate model	YansErrorRate



FIGURE 8. Simulation results for the average registration time under different protocols with optimal configurations.

B. OBSERVATION OF DIFFERENT PROTOCOLS

Fig. 8 depicts a comparison of the proposed HSCT scheme with different protocols to evaluate the registration process at different network sizes. The proposed HSCTopt achieves substantial improvement over all existing protocols by an efficient registration procedure that takes, on average, 64% and 87% less time, compared to the CFTopt and CFT-ATA schemes, respectively. In the experiment, the devices select random value U_R only on initialization, not in every BI. CFT-ATA needs a longer registration time because it has no waiting mode. Therefore, if V_{ACT} reaches its maximum value, it cannot overcome its optimal value. As we can see, both the CAA scheme and the conventional IEEE 802.11ah registration scheme show a rapid increase in registration time for the same fixed step size (Δ). On the other hand, even in a large network with the total number of devices up to 8000, HSCTopt can still maintain better performance than other protocols. In the CAA scheme, it seems that all

registration procedures (including AuthReq/AuthResp and AssocReg/AssocResp) are successfully executed by reserving the channel with the help of the NAV. Although it reserves the channel, since it considers only the CSMA/CA mechanism, massive contention is generated by heavy traffic. Even though the CAC method and the combined authentication/association technique are used to enhance performance, using a fixed Δ makes a big difference in performance efficiency, compared to the HSCTopt. Similarly, IEEE 802.11ah also uses the CAC method, but it deals with contentions twice as well as the CAA scheme with similar limitations. As shown in Fig. 8, IEEE 802.11ah takes a longer time to complete the whole registration process, compared with the HSCT procedure, which creates less contention in the C-slots for AuthRegs, and the rest of the frames are exchanged using the contention-free T-slot. We have taken the optimal value for all the protocols.

C. OBSERVATION OF DIFFERENT ALGORITHMS FOR OPTIMAL ACT VALUE

Figs. 9 and 10 compare the total registration time plotted against the number of devices, with different algorithms for both the HSCT and the CFT schemes, respectively. The optimal algorithm provides the best performance, because it assumes the exact number of devices before starting the registration procedure. The AP can provide the optimal step size (Δ_{opt}) to get the maximum registrations in a BI, based on Eq. (19). Figs. 9 (a) and (b) show the results with 1000 and 8000 devices for the HSCT scheme. In this analysis, the smart-up and smart-down algorithms outperform the up and down algorithms in both small and large networks. The reasons behind are, firstly, the waiting mode initializes V_{ACT} by half of the maximum value instead of the maximum value. The lesser value of V_{ACT} allows a moderate number of devices to send AuthReqs. Therefore, it takes less convergence time to switch from waiting mode to studying mode. Secondly, in studying mode, consideration of both Q_L and S_A allows more devices to select an optimal Δ . Finally, incrementing V_{ACT} in working mode also allows more devices to take both request and response traffic. A small performance gap is observed between the smart-up and smart-down algorithms, which allows more devices in studying mode in the smart-down algorithm than in the smart-up algorithm. In Fig. 9 (b), HSCTopt takes 37.5% less time than HSCT-Up, 30% less time than HSCT-Down, 15% less time than HSCT-Smart-Up, and 12% less time than HSCT-Smart-Down with 8000 devices. Figs. 10 (a) and (b) show the CFT scheme at 1000 and 8000 devices, respectively. In Fig. 10 (b), the registration time under CFTopt is 64% lower than CFT-Up, 53% lower than CFT-Down, 18% lower than CFT-Smart-Up, and 10% lower than CFT-Smart-Down for 8000 devices.

D. OBSERVATION OF SCP AND STP DURATION ON REGISTRATION TIME

Fig. 11 (a) shows the average number of registered devices (per second) at different SCP durations. Since the sum



FIGURE 9. Comparisons of the average registration time of HSCT MAC with different algorithms: (a) the number of devices range from 0 to 1000; (b) the number of devices range from 0 to 8000.



FIGURE 10. Comparisons of the average registration time of CFT MAC with different algorithms: (a) the number of devices range from 0 to 1000; (b) the number of devices range from 0 to 8000.

of D_{SCP} and D_{STP} in T_{BI} is a fixed value $(T_{BI} - T_{BP})$, there is the possibility of an increment/decrement in SCP duration. If the duration of the SCP is increased, then the duration of the STP is decreased, and vice-versa. We can see that the average number of registered devices per second increases with an increased SCP duration, and a long SCP duration allows more AuthReqs; however, it grows linearly up to the highest value, at 41 ms, when the fixed BI is 100 ms, and after that, the performance decreases in a reverse pattern. Although, a long duration for the SCP provides more successful AuthReqs, it reduces the *B* T-slots in the STP duration. Therefore, the management queue of AuthReqs in the AP becomes large, and devices experience a longer delay to complete the registration process.

E. OBSERVATION OF DIFFERENT DISTRIBUTIONS ON REGISTRATION TIME

Fig. 11 (b) compares the registration times when using the uniform and Sift distributions. In the HSCT scheme, at the beginning of each C-slot, unlike conventional CSMA/CA, the access group devices reset the backoff slot. Since the network size is large, the uniform distribution generates high contention at the beginning of the selected C-slot; therefore, the average backoff time and the number of retransmissions will increase for the transmission of frames. On the other hand, the Sift distribution generates the average backoff time and the number of retransmissions at the beginning of every C-slot. It reduces the average backoff time and the number of retransmissions to allow more AuthReqs within the SCP duration. Therefore, the Sift distribution consumes on average 12.5%



FIGURE 11. Simulation results of HSCT MAC with different distributions: (a) average number of registered devices (per second) with different SCP durations and a fixed BI of 100 ms; (b) average registration time of HSCT MAC with Sift and uniform distributions.



FIGURE 12. Simulation results of HSCT MAC: (a) average registration time with different numbers of C-Slots (num_C_Slots); (b) average registration time with 50 interfering devices (traffic rate = 2 Kbps per device).

less time than the uniform distribution in the registration process.

F. OBSERVATION OF THE NUMBER OF CSMA SLOTS (num_C_Slots) ON REGISTRATION TIME

In Fig 12 (a), we observe the importance of the *num_C_Slots* in one SCP, which is one of the major differences from traditional CSMA/CA. In this scheme, the SCP duration is partitioned into multiple mini-CSMA/CA access slots (C-slots). The AP selects the total *num_C_Slots* and assigns these slots with sequential numbering ($C - Slots = 1, ..., num_C_Slots$). Moreover, partition of the

SCP duration creates *sub-access* groups from the access group to limit the number of devices participating in channel contention. We executed the simulation considering different numbers of total CSMA slots, where $num_C_Slots = 1$ means there is no partition in the SCP duration. On the other hand, $num_C_Slots = 2$ makes one partition, which provides two C-slots, and so on. We can see that the division of the SCP duration becomes important for all network sizes. When $num_C_Slots = 3$, 20% less registration time is required compared to $num_C_Slots = 1$ because the traditional CSMA considers all devices in an access group, which generates a massive amount of contention. On the other hand, an optimal number for partitioning the SCP duration is required, where more partitioning reduces contention but increases channel overhead, with more conflict periods and an inter-slot gap, by providing a guard time. Fig. 12 (a) shows that the HSCT with $num_C_Slots = 3$ provides a little bit better performance than $num_C_Slots = 4$. Therefore, $num_C_Slots = 3$ was chosen as the optimal partition for a 100 ms BI in the analysis.

G. OBSERVATION OF SIMULTANEOUS TRANSMISSION

To evaluate the relationship between data traffic and registration time, different amounts of data traffic were considered in our experiments, as depicted in Fig. 12 (b). The network contained 8000 devices, and 50 interfering devices simultaneously generated a total of 100 Kbps, where each device produced 2 Kbps. The devices send PS-poll requests to the AP to access T-slots for data transmission. In T-slot access, priority is given to data transmission over the registration process. In these experiments, each device transmits one packet every X seconds, with $X = R \times L/D$, where R, L, and D are the number of devices, the payload size, and the total data traffic load, respectively. According to Fig. 12 (b), we find the completion of registration takes longer with more data traffic because when the devices demand T-slots for data transmission, there is a reduction in T-slots for the registration process. Therefore, the devices need to wait longer to send the rest of the handshake registration frames through the T-slots.

V. CONCLUSION

In this paper, we proposed a hybrid slotted-CSMA/ CA-TDMA MAC protocol that provides an efficient and scalable registration procedure for machine-to-machine communications by large numbers of IoT devices (up to 8000). We proposed the use of multiple C-slots to make the network compatible, and to avoid congestion in the presence of massive numbers of M2M devices. Under HSCT, several mechanisms are used to efficiently control the massive amount of contention: i) AuthReqs are processed by an efficient medium access scheme that is achieved by dividing the slotted-CSMA/CA period into multiple C-slots; ii) adaptive adjustment of the SCP and STP durations minimizes any waste of channel bandwidth; iii) contention is mitigated by providing a contention-free T-slot for the AuthResp and AssocReq/AssocResp frames (the combined scheme of contention-based C-slot access for AuthRegs and contentionfree TDMA access for the remaining message exchanges enhances the overall performance of the proposed HSCT protocol); iv) Sift geometric probability distribution is used to minimize the collision probability among contending devices during the SCP, increasing the devices' transmission probability; and v) the CAC method is used to make groups improve performance against the massive number of contentions by using two modified algorithms: smart-up and smart-down. We provided a closed-form analytical model for the number of AuthRegs in the C-slots. In addition, the performance of the proposed HSCT was analyzed using NS-3 network simulations under IEEE 802.11ah with modifications for the proposed scheme. The proposed HSCT was compared with the IEEE 802.11ah standard and CFT. From the simulation results, the robustness and scalability of the HSCT MAC were confirmed by its ability to complete registration procedures, on average, in 64% and 87% less time, compared to the existing CFTopt and CFT-ATA schemes. Moreover, we evaluated the optimal setting of parameters for the authentication control threshold to maximize the number of devices sending an AuthReq message.

REFERENCES

- J. Gubbi, R. Buyya, S. Marusic, and M. Palaniswami, "Internet of Things (IoT): A vision, architectural elements, and future directions," *Future Generat. Comput. Syst.*, vol. 29, no. 7, pp. 1645–1660, 2013.
- [2] D. Lund, C. MacGillivray, V. Turner, and M. Morales, "Worldwide and regional Internet of Things (IoT) 2014–2020 forecast: A virtuous circle of proven value and demand," Int. Data Corp., Framingham, MA, USA, Tech. Rep. #248451, May 2014, pp. 1–27.
- [3] IEEE Standard for Information Technology—Telecommunications and Information Exchange Between Systems—Local and Metropolitan Area Networks—Specific Requirements—Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications Amendment 2: Sub 1 GHz License Exempt Operation, IEEE Standard 802.11ah-2016 (Amendment to IEEE Std 802.11-2016, as amended by IEEE Std 802.11ai-2016), May 2017, pp. 1–594, doi: 10.1109/IEEESTD.2017.7920364.
- [4] T. Adame, A. Bel, B. Bellalta, J. Barcelo, and M. Oliver, "IEEE 802.11AH: The WiFi approach for M2M communications," *IEEE Wireless Commun.*, vol. 21, no. 6, pp. 144–152, Dec. 2014, doi: 10.1109/MWC.2014.7000982.
- [5] H. Wang. (2012). Supporting Authentication/Association for Large Number of Devices. [Online]. Available: http://mentor.ieee.org/802.11/ dcn/12/11-12-0112-04-00ahsupporting-of-the-authentication-associationfor-large-number-of-devices.pptx
- [6] T.-H. Hsu and P.-Y. Yen, "Adaptive time division multiple access-based medium access control protocol for energy conserving and data transmission in wireless sensor networks," *IET Commun.*, vol. 5, no. 18, pp. 2662–2672, Dec. 2011.
- [7] P. Sthapit, S. Subedi, G. R. Kwon, and J. Y. Pyun, "Performance analysis of association procedure in IEEE 802.11ah," in *Proc. 10th Int. Conf. Syst. Netw. Commun.*, 2015, pp. 70–73.
- [8] D. Bankov, E. Khorov, and A. Lyakhov, "The study of the centralized control method to hasten link set-up in IEEE 802.11ah networks," in *Proc.* 21th Eur. Wireless Conf., May 2015, pp. 1–6.
- [9] D. Bankov, E. Khorov, A. Lyakhov, and E. Stepanova, "Fast centralized authentication in Wi-Fi HaLow networks," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Paris, France, May 2017, pp. 1–6, doi: 10.1109/ICC. 2017.7996510.
- [10] D. Bankov, E. Khorov, and A. Lyakhov, "The study of the distributed control method to hasten link set-up in IEEE 802.11ah networks," in *Proc. 15th Int. Symp. IEEE Problems Redundancy Inf. Control Syst. (REDUN-DANCY)*, Sep. 2016, pp. 13–17.
- [11] N. Shahin, L. Tann, and Y.-T. Kim, "Enhanced registration procedure with NAV for mitigated contentions in M2M communications," in *Proc. 18th Asia–Pacific Netw. Oper. Manage. Symp. (APNOMS)*, Kanazawa, Japan, Oct. 2016, pp. 1–6.
- [12] R. Ali, S. W. Kim, B. S. Kim, and Y. Park, "Design of MAC layer resource allocation schemes for IEEE 802.11ax: Future directions," *IETE Tech. Rev.*, vol. 35, no. 1, pp. 28–52, Nov. 2016, doi: 10.1080/ 02564602.2016.1242387.
- [13] Y. Liu, C. Yuen, X. Cao, N. U. Hassan, and J. Chen, "Design of a scalable hybrid MAC protocol for heterogeneous M2M networks," *IEEE Internet Things J.*, vol. 1, no. 1, pp. 99–111, Feb. 2014, doi: 10.1109/ JIOT.2014.2310425.
- [14] Y. C. Tay, K. Jamieson, and H. Balakrishnan, "Collision-minimizing CSMA and its applications to wireless sensor networks," *IEEE J. Sel. Areas Commun.*, vol. 22, no. 6, pp. 1048–1057, Aug. 2004.
- [15] R. de Vegt. (2011). Potential Compromise for 802.11ah Use Case Document. [Online]. Available: http://mentor.ieee.org/802.11/dcn/11/11-11-0457-00-00ah-potentialcompromise-of-802-11ah-use-casedocument.pptx>R

- [16] R. Zhang, L. Cai, and J. Pan, "Performance analysis of reservation and contention-based hybrid MAC for wireless networks," in *Proc. IEEE ICC*, May 2010, pp. 1–5.
- [17] High Rate Ultra Wideband PHY and MAC Standard, ECMA International Standard ECMA-368, Dec. 2005. [Online]. Available: http://www.ecmainternational.org/publications/standards/Ecma-368.htm
- [18] R. Zhang, L. Cai, and J. Pan, "Performance study of hybrid MAC using soft reservation for wireless networks," in *Proc. IEEE Int. Conf. Commun.* (*ICC*), Kyoto, Japan, Jun. 2011, pp. 1–5, doi: 10.1109/icc.2011.5963319.
- [19] R. Ruby and J. Pan, "Video streaming with PCA and hard vs soft DRP," in *Proc. IEEE GLOBECOM*, Miami, FL, USA, Dec. 2010, pp. 1–6.
- [20] Q. Ye and W. Zhuang, "Distributed and adaptive medium access control for Internet-of-Things-enabled mobile networks," *IEEE Internet Things J.*, vol. 4, no. 2, pp. 446–460, Apr. 2017, doi: 10.1109/JIOT.2016.2566659.
- [21] IEEE Standard for Information Technology Telecommunications and Information Exchange Between Systems Local and Metropolitan Area Networks Specific Requirements Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications, IEEE Standard 802.11-2012, Mar. 2012.
- [22] L. Tian, S. Deronne, S. Latré, and J. Famaey, "Implementation and validation of an IEEE 802.11ah module for ns-3," in *Proc. Workshop ns-3. ACM*, 2016, pp. 49–56.
- [23] The Network Simulator-ns-3. [Online]. Available: https://www.nsnam.org/
- [24] A. Hazmi, J. Rinne, and M. Valkama, "Feasibility study of IEEE 802.11ah radio technology for IoT and M2M use cases," in *Proc. IEEE Globecom Workshops*, Dec. 2012, pp. 1687–1692.
- [25] M. Park, "IEEE 802.11ah: Sub-1-GHz license-exempt operation for the Internet of Things," *IEEE Commun. Mag.*, vol. 53, no. 9, pp. 145–151, Sep. 2015, doi: 10.1109/MCOM.2015.7263359.
- [26] M. Park, "IEEE 802.11ah: Energy efficient MAC protocols for long range wireless LAN," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2014, pp. 2388–2393.



NURULLAH SHAHIN received the B.Sc. and M.Sc. degrees from the Department of Information and Communication Engineering, Islamic University, Kushtia, Bangladesh, in 2009 and 2010, respectively. He is currently pursuing the combined M.Sc. and Ph.D. degree with the Advanced Networking Technology Lab, Department of Information and Communication Engineering, Yeungnam University, Gyeongsan, South Korea. He is currently a Maintenance Engineer

with the IT Operation and Communication Department, Bangladesh Bank (Central Bank of Bangladesh), Bangladesh. His research interests include dense wireless networks, vehicular ad-hoc networks, and resource allocation in wireless networks.



RASHID ALI received the B.Sc. degree in information technology from Gomal University, Pakistan, in 2007, and the M.Sc. degree, under the supervision of Dr. S. Belenki, in computer science (advanced network design) in 2010, and the M.Sc. degree in informatics, under the supervision of Dr. M. Spante, from University West, Sweden, in 2013. He is currently pursuing the Ph.D. degree with the Wireless Information Networking Lab, Department of Information and Communication

Engineering, Yeungnam University, South Korea. From 2007 to 2009, he was with Wateen Telecom Pvt., Ltd., Pakistan, as a WiMAX Engineer with the Operations Research Department. From 2013 to 2014, he was with the COMSATS Institute of Information Technology, Vehari, Pakistan, as a Lecturer. His research interests include the enhancement of efficiency and reliability in future WLANs, modeling and analyzing the stochastic process of MAC layer resource allocation in future WLANs, and the Internet of Things



YOUNG-TAK KIM (S'84–M'90) received the Ph.D. degree from KAIST in 1990. He joined Korea Telecom (KT) in 1990, where he researched and developed the ATM MAN switching system and related network operations and management technologies for broadband networking. He participated in the standardization activities of the ITU-T Study Group 13 as the Representative of KT. In 2001, 2008, and 2015, he was a Visiting Scholar with the National Institute of Standards

and Technology (NIST), USA, where he joined in the design and implementation of the NIST GMPLS simulator (GLASS) and the Next Generation Routing Architecture (NGRA) Project. He is currently a Professor with the Department of Information and Communication Engineering, College of Engineering, Yeungnam University, South Korea. His research interests include QoS-guaranteed traffic engineering in future Internet, cloud computing, QoS-aware network operating systems, OpenFlow, QoS-aware seamless secure mobility, software-defined networking/network function virtualization, and related network operations and management. He is a member of the IEEE Communications Society, KICS, KISS, KIPS, and the Korea Multimedia Society. He was the Technical Program Chair of the IEEE ComSoc CNOM for 2007–2008, a TPC Co-Chair of the IFIP/IEEE IM2009, a General Chair of APNOMS2009, and a TPC Co-Chair of the IEEE/IFIP NOMS 2018.

...