Towards An Adaptive Time-Triggered Protocol in Wireless Networks

Jin Zhang*, Fan Liang*, Wei Yu*, David Griffith†, Wenqi Guo†, and Avi Gopstein†
*Towson University, USA
Emails: {jzhang13,fliang1}@students.towson.edu, wyu@towson.edu
†National Institute of Standards and Technology (NIST), USA
Emails: {david.griffith, wenqi.guo, avi.gopstein}@nist.gov

***Abstract*—Time sensitive activities occur extensively in industrial Internet of Things (IoT) environments. Classical time-triggered protocols in wired networks, such as Ethernet, have been proven effective in supporting real-time and safety-critical communications. Along with the increasing use of wireless devices, designing wireless network protocols that handle the delivery of time-sensitive activities becomes an important issue. Unlike the Time-Triggered Ethernet (TTE) protocol designed for wired networks, a time-triggered protocol tailored to wireless networks faces a number of barriers, such as unstable data transmission under high radio interference, high bit error rate, and limited computation and storage capabilities on IoT devices. In this paper, we introduce an Adaptive Time-Triggered Protocol (ATTP) for wireless networks. In particular, we design a dynamic repetition scheme that transmits replicated copies of one time-triggered data packet to ensure the reliability and low latency of data transmission. To minimize the overhead raised by the replication of packets, we dynamically select the necessary number of replicated copies according to the instantaneous quality of the connection link. To effectively utilize wireless resources, we also design a multi-class traffic scheduling scheme to handle three traffic types, time-triggered (TT) traffic that requires an absolute latency guarantee, event-triggered (ET) traffic that requires a relatively low latency guarantee, and best-effort (BE) traffic without any latency guarantee requirement. Via an extensive simulation study, we validate the effectiveness of our approach, achieving expected performance for time-triggered traffic with respect to latency, packet loss ratio, and overhead reduction.**

***Keywords*—Wireless networks, Time-Triggered Protocols, Time Sensitive Traffic, Real-time Communication, Self-adaptive system, IoT, LTE/LTE-A**

I. Introduction

The Internet of Things (IoT) enables the connection of a variety of smart devices so that different physical objects can be monitored and controlled, leading to the success of numerous smart-world systems. IoT has been widely applied to diverse fields, including the smart grid, smart manufacturing, smart transportation, smart health, smart farming, smart home, and smart cities, among others [1], [2], [3], [4]. Of these applications, some demand strict quality of service (QoS) for data transmission. For example, controlling a manufacturing system or powering electricity transmission in the smart grid commonly require real-time data delivery, either from critical sensors to a control center, or from a control center to critical actuators. To handle real-time data traffic, time-triggered protocols (TTP) were proposed [5].

One representative implementation of TTP, Time-Triggered Ethernet (TTE), is an extended network technology that builds the TTP protocol on top of IEEE 802.3 [6]. TTE utilizes dedicated time slots to guarantee the low latency of real-time traffic over Ethernet. TTE can schedule and optimize network resources so that desirable performance for real-time traffic delivery over Ethernet can be achieved. Therefore, it can satisfy the performance requirements of critical real-time applications, and has thus been widely adopted as a viable solution in wired networks, supporting real-time and safety-critical systems in areas such as aerospace, automobiles, and manufacturing, among others.

Nonetheless, along with the rapidly increasing volume of wireless devices in IoT, massive volumes of data need to be delivered through wireless networks with real-time latency requirements. Thus, similar to Ethernet delivery, it is necessary to design a protocol to ensure the transmission latency of real-time traffic over wireless networks. Nonetheless, compared to the stable transmission environment of wired networks, there are a number of challenges to support real-time data transmission in wireless networks. First, the wireless network is unstable, as it is more susceptible to interference. The open radio communication environment of wireless networks incurs strong radio interference, which can pose a high bit error rate. In addition, the transmission latency can be largely affected by the quality of wireless communication channels. Second, IoT devices have a number of resource constraints, such as limited computing capabilities and battery resources. Therefore, the protocols for supporting real-time communications need to be highly efficient with low computing overhead. Because of the outlined challenges in wireless networks, the existing TTP protocols cannot be directly applied to wireless networks.

To resolve these issues, in this paper, we make the following contributions.

- **Dynamic Replication:** We propose an Adaptive Time-Triggered Protocol (ATTP) for wireless networks and design the dynamic repetition scheme. To achieve reliability in data transmission, multiple replicated copies of the same data packet are transmitted simultaneously. As replicating data packets consumes extra network bandwidth resources, we develop an adaptive replication scheme that dynamically selects the necessary number of replicated packets to balance the reliability and latency of data transmission and its overhead to the network. A communication channel with higher packet loss ratio is given a higher repetition number. By doing this, our approach achieves low latency and high data reliability for time-triggered traffic with marginal overhead to the network.

- **Multi-class Traffic Scheduling:** In addition to the time-triggered (TT) traffic that requires highly reliable real-time data delivery, we consider two other non-real-time traffic types, which are event-triggered (ET) traffic and best-effort (BE) Traffic. To satisfy the performance requirements of different traffic types, the pre-scheduled transmission time slots (i.e., TT traffic time slots) is reserved to transmit the TT traffic, while ET and BE packets are transmitted during the non-TT traffic time slots. Moreover, ET traffic takes precedence over BE traffic. Consequently, our approach provides low latency for TT traffic; meanwhile, it improves the efficiency of network resource utilization by handling both ET and BE traffic.
- **Performance Evaluation:** We implement our approach in NS-3, a well-known open-source network simulator for networking research [7][1]. We design comprehensive experiments to evaluate the effectiveness of our approach as compared with other approaches. The results of our experiments demonstrate that TT traffic in our approach attains lower latency, smaller packet loss ratio, and lower overhead, compared with the other approaches.

The remainder of this paper is organized as follows: In Section II, we review existing research efforts. In Section III, we provide preliminaries on time-triggered protocols and channel quality. In Section IV, we introduce our approach in detail, including problem space, dynamic replication scheme, and multi-class traffic scheduling scheme. In Section V, we present our evaluation methodology and results. Finally, we conclude the paper in Section VI.

II. Related Works

To support time-sensitive data transmission, TTPs have been designed to support real-time applications, critical safety communications, and industrial control systems. In particular, TTE was proposed and standardized based on wired Ethernet networks [8]. Considering the requirement of highly time sensitive applications, a body of research efforts have been conducted on network performance for time sensitive applications, especially on wireless networks [9], [10], [11], [12], [13], [14], [15], [16], [17], [18], [19], [20]. Existing research works focus on a variety of perspectives, including TT protocol design and development, improvement of time sensitive network performance, and constructing time-triggered systems on wireless networks.

[1]Certain commercial equipment, instruments, or materials are identified in this paper in order to specify the experimental procedure adequately. Such identification is not intended to imply recommendation or endorsement by the National Institute of Standards and Technology, nor is it intended to imply that the materials or equipment identified are necessarily the best available for the purpose.

A number of research efforts have also been dedicated to designing and improving TTP [9], [10], [12]. For example, Kopetz *et al.* [9] improved the topology of TTP networks for a fault-tolerant real-time systems. They used predictable message transmission, message acknowledgment and replicated hardware (i.e., replicated nodes and replicated communication channels) to meet the QoS requirement of high dependability and guaranteed timeliness. Steinbach *et al.* [10] validated that a new system based on time-triggered Ethernet could be utilized to replace FlexRay in existing in-vehicle bus-systems. Note that the in-vehicle system using switch technology has the benefit of bandwidth utilization over shared bus technology, when group communication is used. Likewise, Steiner [12] designed a customized satisfiability modulo theories (SMT)-based algorithm to solve problem of TT transmission scheduling. The experimental results showed that the developed SMT-based algorithm could solve the complicated multi-hop TT communication scheduling problem and produce up to ninety percent maximum utilization on a communication link.

Related to the development of time sensitive networks, some research efforts have been conducted to improve network performance [13], [14]. For example, Shan *et al.* [13] developed a new heuristic called the top-down algorithm for real-time data gathering in wireless sensor networks. The designed algorithm could prolong the network lifetime by constructing the routing tree layer by layer using the network flow model. Likewise, Shrestha *et al.* [14] designed an enhanced precision time protocol (PTP) by incorporating a clock drift factor into clock synchronization, which is a critical component for time sensitive industrial wireless sensor networks deployed for critical control and automation applications. The designed protocol could improve the accuracy of clock offset estimation for conventional PTP schemes.

Regarding the design of TTP on wireless networks, there have been a number of efforts [15], [16], [18]. For example, Bartolomeu *et al.* [18], [15] explored the architecture and protocol operation of a wireless flexible time triggered (WFTT) technique, which is used for IoT applications with stringent timing requirements. Via analyzing the specifications, implementation, and performance of WFTT with bandjacking technique, the developed WFTT protocol was confirmed to have the capability of supporting real-time communications even in high interference wireless networks such as WiFi. Likewise, Jacob *et al.* [16] designed a time-triggered wireless network for cyber-physical systems, which attempts to minimize communication energy consumption and offer end-to-end timing predictability, runtime adaptability, reliability, and low latency.

Unlike existing research efforts, in this paper we propose our own approach to effectively deliver time sensitive data in wireless networks. In our approach, we design a dynamic repetition scheme so that real-time and reliable data transmission of TT messages can be supported while overhead to the network can be minimized. Additionally, we design a multi-class traffic scheduling scheme that other non-real-time traffic types can be scheduled and delivered together with real-time

TT traffic, improving network utilization.

III. Preliminary

In this section, we introduce the concepts of TTP and channel quality.

A. Time-Triggered Protocol (TTP)

One distinguishing characteristic of time-triggered systems is that all TT packet transmissions are pre-scheduled globally by all the TTP nodes. Thus, time synchronization on all TTP nodes is necessary. The time synchronization process establishes an initial agreement on all TTP network components. Then, the TTP operates cyclically following a schedule table, which determines the time when a message can be transmitted by a node. Thus, the provision of a system-wide global time base with good precision and accuracy is a key prerequisite for the design of time-triggered systems. There are distinct synchronization schemes depending on the system models and topologies, including client/server and completely distributed synchronization.

Data communication in TTP is organized in time-division multiple access rounds. TT nodes transmit data packets periodically. The TT traffic scheduling considers two dimensions, the time slots and bandwidth (resources). Each TT node obtains transmitting time slots as shown in Fig. 1. On the bandwidth (resource) dimension, several nodes may transmit packets on the same time slot to utilize bandwidth efficiently. Here, $TTm1$, $TTm2$, $TTm3$, and others represent the packets that are sent by TT nodes 1, 2, 3, and others. Then, after a TT cycle, TT nodes start the next round to send out data. Note that, as shown in Fig. 1, if TT nodes do not have data to send, their assigned resources go unused. In addition, BE nodes are able to transmit only in the resources that are left available by ET nodes that do not have data to send.

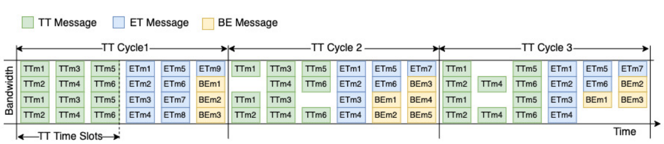

Fig. 1. Traffic Scheduling

B. Channel Quality

The quality of communication channels is a key indicator for TTP wireless communications. Usually, the wireless channel condition is represented by signal-to-noise ratio (SNR), signal-to-interference plus noise ratio (SINR), signal-to-noise plus distortion ratio (SNDR). Among these three metrics, the main criterion to determine the quality of the channel condition is SNR. High SNR indicates a high channel condition quality. The value of SNR varies with different channel models (referred to transmission path loss), fading speed, fixed or moving transmitter/receiver, and the number of transmitter/receiver antennas. Signal y received by a receiver antenna is related to transmitter signal x and noise n, formalized via $y = h * x + n$. Here, h is the vector of correlations between transmitter and receiver antenna and y, x, and n are vectors based on the number of transmitters and receivers.

In our case, we assume the environment white noise follows a Gaussian distribution. Also, there is no correlation between white noise of different environments. We denote P as total transmission power and w^H as the conjugate transpose of the weight factor vector on receiver side antenna. According to [21], the SNR can be represented by $SNR = \frac{\textit{Signal Power}}{\textit{Noise Power}} = |w^H h|^2 \frac{P}{{\sigma_n}^2}$. Here, σ_n is represented by ${\sigma_n}^2 = \mathbb{E}\{\sum_{i=1}^{r} |n_i|^2\}$, where σ_n is a critical and dynamic factor for SNR. Its value expresses the channel condition of wireless link. A high value for σ_n means a poor condition for decoding received signal, leading to higher bit error rate.

IV. Our Approach

In this section, we introduce our approach in detail. Particularly, we first provide an overview of the problem space and network topology. We then introduce the adaptive repetition scheme. Finally, we present the multi-class scheduling scheme to handle the delivery of ET and BE non-real-time traffic along with real-time TT traffic.

A. Overview

Fig. 2 describes the problem space for time-triggered systems, which consists of three dimensions (i.e., traffic types, transmission media, and QoS requirements) that define eighteen sectors. The three solid blue sectors in the figure indicates our research focus in this study, which involves leveraging the principle of TTP in wireless networks so that the performance requirements for real-time applications can be satisfied over unstable wireless communication channels. In detail, we utilize TTP to reduce transmission latency and guarantee real-time performance for time-triggered applications on wireless networks. In our design, we adopt a dynamic repetition scheme based on the wireless channel quality, instead of the typical acknowledgment and re-transmission scheme that has been widely used to ensure the reliability of data transmission. Considering both real-time and non-real-time traffic in real-world network environments, we also include ET and BE traffic as two types of non-real-time traffic that must be accounted for, and design a multi-class traffic scheduling scheme so that network resources can be used effectively.

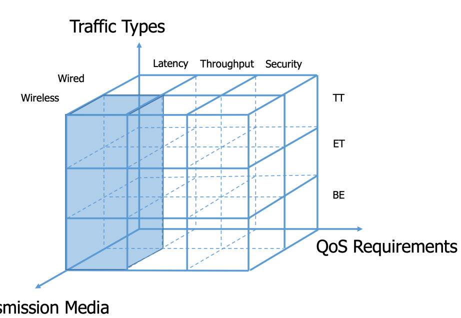

Fig. 2. Problem Space of Network Applications

In our study, we consider a cascade star network in an IoT environment. All IoT devices connect with a controller via wireless communication. All controllers connect to the management center via a wired network as shown in Fig. 3. Controllers process the time synchronization and TT traffic transmission scheduling, which are the key components in ATTP. The client-server synchronization scheme is used in ATTP. The Management Center (as a server node) distributes the synchronization signal to all controllers, and controllers forward the synchronization beacon to IoT devices (the client nodes) connected with them. Then, the controllers allocate time slots and create the TT traffic scheduling table.

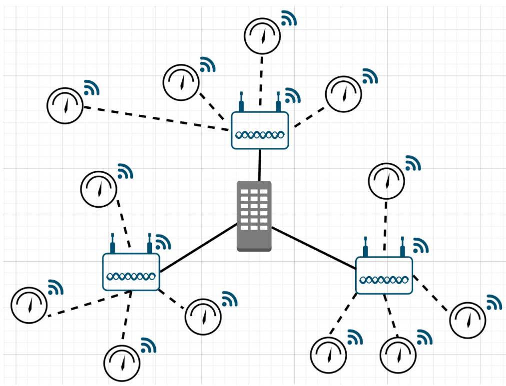

Fig. 3. Topology of ATTP

B. Adaptive Repetition Scheme

In non-TTP wireless networks, such as IEEE 802.11 and IEEE 802.15.4, the acknowledgment and re-transmission scheme has been adopted to provide transmission reliability. In this way, the transmitter keeps packets in the buffer until receiving the acknowledge message back from the receiver. The acknowledgment and re-transmission scheme incurs both waiting and re-transmission time. Thus, it is not appropriate for supporting many real-time applications.

Instead, in our scheme, replication causes TT nodes to transmit several replicated copies of one TT packet to provide reliability and latency guarantees in data transmission. The repetition improves the reliability of critical data transmission. As shown in Fig. 1, if one copy of TT message 2 is lost in TT cycle 2, the receiver will still receive the message due to another duplicate arriving on time. The same situation occurs with TT message 5. Likewise, if one copy of message 5 is lost, the receiver can still receive another copy of the message to ensure message 5 is delivered.

As shown in Fig. 1, these replicated TT messages will be transmitted within the same time slot. Even if one packet is lost during transmission, the other copies can still be delivered to the receiver, which avoids re-transmission and improves latency. Compared to the acknowledge and re-transmission scheme, our approach can significantly reduce latency and achieve reliability and low latency in data transmission. In addition, compared to the acknowledgment and retransmission scheme, repetition requires less computation and storage resources, as it does not keep the packets in the buffer to wait for acknowledgments. This makes it more suitable for IoT devices with limited resources and low computation capabilities.

Nonetheless, transmitting replicated TT packets will introduce more overhead to the network as replicated packets for the same data are transmitted over the network. To address this issue, we introduce a mechanism to dynamically select the number of replicated copies of TT data packets according to channel quality measured by SNR (described in Section III-B). In a wireless industrial control scenario, IoT devices report the instantaneous channel quality to the controller and the controller adjusts the replication number for individual IoT nodes depending on the report. To make our algorithm robust, both upper and lower thresholds for determining channel quality is established. Algorithm 1 shows the procedure of adaptive replication. In the algorithm, r represents the replication number, σ represents the channel quality indicator (CQI) based on SNR, λ represents the lower CQI threshold to determine whether r should be increased, and θ represents the upper CQI threshold to determine whether r should be reduced. Additionally, c_d is the counter for CQI being lower than λ, so that r should be increased, and c_i is the counter for CQI being higher than θ, indicating that r should be reduced. The initial value for c_d and c_i are 0. Also, c_{th} is the counter threshold to check the accumulated counters c_d and c_i, t is the current time, and *TP* is the time window for updating counters c_d and c_i.

Algorithm 1 Adaptive Replication

Require: Replication number r, Channel Quality Indicator σ, Lower Threshold λ, Upper Threshold θ, c_d and c_i are counters (starting from 0), and c_{th} is the counter threshold

while every TT cycle **do**
  if $t >= TP$ **then**
    if $c_d > c_{th}$ **then**
      Increment r and update time slot schedule table
    end if
    if $c_i > c_{th}$ **then**
      Decrement r and update time slots schedule table
    end if
  else
    if $\sigma < \lambda$ **then**
      Increment c_d
    end if
    if $\sigma > \theta$ **then**
      Increment c_i
    end if
  end if
end while

We now analyze the average latency of data packets transmitted via our approach. Denote t as the transmission time for one packet transmitted from TT nodes to the controller, $r \in \mathbb{N}$ as the replication number, and e as the bit error rate. For one data packet, r copies of the packet are transmitted within the same time slot. $T_t = \{t_1, t_2, \ldots, t_r\}$ is the set of all transmission times for these packet copies. If the receiver gets the first copy correctly, the transmission time for this packet will be $t_{min_1} = \min\{t | t \in T_t\}$. If an error occurs, the transmission time of the second copy is the actual transmission delay and can be represented by $t_{min_2} = \min\{t | t \in T_t - t_{\min_1}\}$. If there are still some errors in the second copy, the transmission time of the third copy should be the actual transmission delay, and so on. For $1 < i < r$, the average transmission latency t of a packet using repetition r is $t = \sum_{i=1}^{r} t_{\min_i} e^{i-1}(1-e)$, where e is the probability of a packet error, and we assume that packet error events are independent of each other.

Consider the traditional acknowledgment and re-transmission scheme to ensure the reliability of data transmission. Denote t_r as the transmission time of a packet from TT node to the controller and t_a as the acknowledgment signal transmission time from the controller to TT nodes. The average packet transmission time in non-TT traffic, assuming a single retransmission, is $t = t_r(1-e) + (t_a + 2t_r)e(1-e)$. Comparing the average transmission delay in the acknowledgement/re-transmission scheme to our replication approach, it is obvious that our replication approach achieves improved latency performance on data transmission.

In addition, denote $l_{avg}^{ATTP} = \frac{1}{n}\sum_{i=1}^{n} l^{r_i}$ as the packet loss ratio in our approach. Here, denote replication number for node i as r_i, $n \in \mathbb{N}$ as the number of TT nodes, and l as the packet loss ratio of data transmission. Comparing to the $l_{avg} = \frac{1}{n}\sum_{i=1}^{n} l$ in the non-TTP (i.e., our approach is not used, meaning traditional acknowledgement/re-transmission scheme). We can see the packet loss ratio in our approach is much smaller than that in the non-TTP. Thus, the packet loss ratio in our TT approach will be reduced due to the dynamic repetition scheme.

C. Multi-class Traffic Scheduling Scheme

In addition to handling TT traffic, our ATTP considers two additional types of traffic, ET and BE traffic for the effective use of wireless bandwidth resources. ET and BE traffic will be transmitted within the non-TT period. Also, ET traffic has higher priority than BE traffic. The ATTP only transmits the BE traffic when there is no TT or ET traffic occupied the network bandwidth resource. Compared to the real-time requirement of TT traffic, ET traffic provides high latency guarantee, and BE traffic has no latency guarantee.

The TT packets are transmitted in the pre-scheduled time slots in the TT traffic schedule table. Since the TT traffic has the highest priority, the time slots cannot be preempted by other types of traffic. Moreover, even when there is no TT message in the TT time slots, the ET and BE traffic are not allowed to be transmitted in the TT traffic slots. For example, at TT Cycles 3 in Fig. 1, the time slots reserved for a TT node that transmits TT message 5 are not occupied, but no ET or BE traffic can use the resource.

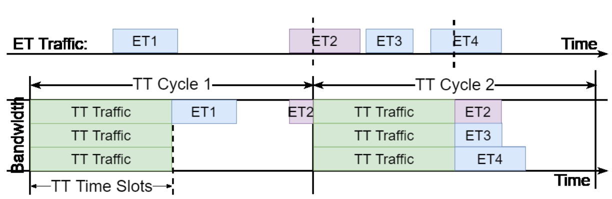

Fig. 4. Event Traffic Scheduling

In the following, we introduce the details of scheduling ET and BE traffic in out ATTP scheme.

- **Event-Triggered Traffic:** In the wireless network system, the ET traffic is the typical traffic which is triggered by specific events. The nodes send the ET packets when the event occurs. Nonetheless, ET traffic has lower priority than TT traffic, the nodes hold ET traffic to avoid the conflicts with TT packet transmission in order to guarantee low latency for TT traffic. In Fig. 4, traffic *ET 1* is generated during TT time slots. Then, it starts transmitting right after the TT traffic period. If traffic *ET 2* is generated before the TT time slot, it will be transmitted instantly and then stopped until the beginning of next TT time slot. After the TT time slot ends, it will be resumed transmission. Similar to *ET 1*, *ET 3* and *ET 4* will be held to avoid transmission during TT time slots.
- **Best-Effort Traffic:** The BE traffic has the lowest priority among all three types of traffic. It utilizes the remaining network resources to transmit data. Specifically, BE nodes scan the TT scheduling table and hold BE traffic to avoid conflicting with TT traffic. In the non-TT time slot, the BE nodes keep monitoring the network resources. They only transmit BE packets when the resources are not fully occupied by ET traffic. If there is no available network resource, as shown in Fig. 5, *BE 1* conflicts with TT traffic, and is held until attempting to re-transmit after a random back-off time. Nonetheless, it still conflicts with ET traffic, as there is no available bandwidth resource left for the second transmission of *BE 1*. Thus, *BE 1* waits for another random back-off time and then gets network resources for its third transmission attempt. *BE 2* starts its data transmission outside the TT time slots, but it is preempted by TT traffic and ET traffic. The remaining traffic of *BE 2* is transmitted after random back-off time. BE traffic has no guarantee of data transmission performance.

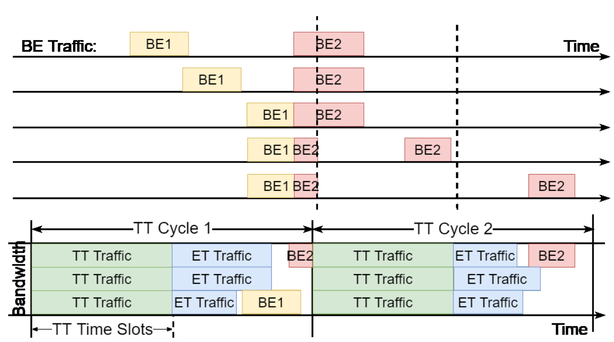

Fig. 5. Best Effort Traffic Scheduling

V. Performance Evaluation

In this section, we first present the evaluation methodology and then describe the evaluation results.

A. Methodology

Environment: To evaluate the proposed ATTP strategy, we use ns-3 [7] as the simulation environment. In detail, we deploy the ns-3 (ns-allinone-3.29) on Ubuntu 16.04. The ns-3 (ns-allinone-3.29) contains both wired and wireless network modules, including WiFi, MANET, LoRaWAN, and LTE. The ns-3 tool allows researchers to create their own protocols and applications in addition to the existing ones. The flexibility of ns-3 to integrate new modules assists users to build complex networks that incorporate novel technologies. We build ATTP on the Long-term Evolution (LTE)/LTE-Advanced (LTE-A) network as an example to demonstrate our design. Note that LTE/LTE-A is a viable wireless network infrastructure for IoT as it provides broadband wireless connections and wide-area coverage. In the LTE network, we adopt eNodeB (eNB) as controllers for our approach. The TT time slots and resource distribution are located in eNB. LTE groups channel resources into resource blocks (RBs). An RB contains two dimensions: (i) frequency domain, and (ii) time domain. One RB consists of 180 kHz in the frequency domain and 0.5 ms (1 time slot) in the time domain. The minimum unit of resource distribution in LTE/LTE-A is a pair of time slots (referred to as one subframe) which consists of two consecutive Resource Blocks (RBs). Radio resource allocation is executed every 1 ms, which is defined as transmission time interval (TTI). The eNB collects the information of connected TT nodes and channel quality and then plans the scheduling table for TT traffic.

Scenarios: To evaluate our proposed ATTP, we design four experimental scenarios: (i) We apply the TTP protocol to the wireless network directly. TT nodes send the data on time slots by following the pre-defined TT traffic scheduling table. TT traffic has the highest priority. ET and BE traffic send data only on the non-TT traffic time intervals. (ii) Based on the TTP protocol, we implement a Repetition Time-Triggered Protocol (RTTP) that enables the repetition strategy to allow the TTP protocol to replicate individual data packets (each data packet is replicated as two copies in our RTTP protocol). (iii) We implement our proposed ATTP approach, which adopts the adaptive replication scheme to transmit TT packets via dynamically adjusting the number of replicated packets according to the channel quality. In the experiment, we adopt two as the repetition number for poor channels. (iv) We evaluate non-TTP as a baseline, which does not consider the priorities of TT, ET, and BE traffic.

Metrics: In our experiment, the following four key metrics are defined and employed to evaluate the performance of the aforementioned scenarios over LTE/LTE-A networks: (i) *Throughput* is defined as the average number of bits per second that the eNB receives from one node during the uplink evaluation of the LTE network. This metric shows the average received bits in unit time from one node on the packet data convergence protocol (PDCP) layer. (ii) *Average Latency* refers to the average delay starting from a time when IoT devices try to send data packets for the first time, to the time when the controller receives them. For TT nodes, latency is the time interval starting from the moment when a packet is sent out from the device's PDCP layer to the moment when the controller's PDCP layer receives the packet. Latency for ET traffic and BE traffic is not only the transmission time from sender to receiver like TT traffic delay, but also considers the back-off time when ET traffic conflicts with TT traffic or the BE traffic conflicts with TT and ET traffics. (iii) *Packet Loss Ratio (PLR)* is the ratio between the amount of actually received data by the controller and the amount of total transmitted data sent from IoT devices on the PDCP layer. (iv) *Goodput* is defined as the amount of bits per second that the controller receives from TT nodes after removing duplicated data packets. The data is collected from the PDCP layer as well.

Simulation Setups: We now describe the setup of our experiment. We simulate the four scenarios in ns-3. Our simulation is built on the LTE wireless networks to demonstrate our idea. In our experiment, we build an eNB as a controller and all UEs (i.e., IoT devices) dispersed around it following a uniform distribution within its coverage. According to the predictable feature of TT traffic, we deploy 20 TT nodes and pre-allocate fixed time slots for them. In the process of simulation, we increase the number of ET nodes and BE nodes to simulate the increase in traffic burden on the wireless network. The time of generating ET and BE traffic follows the uniform random function. All parameters of our simulation over LTE/LTE-A network in ns-3 are listed in Table I.

We position all UE nodes within the eNB's coverage. The bandwidth of the eNB is 5 MHz. The eNB's transmission power is set as 46 dBm. We adopt CAT-0 category on the UE end, which is a new UE category designed for IoT and machine-to-machine networks in LTE Release 12. According to 3GPP TS 36.101, the UE transmission power is set to 23 dBm. The coverage distance is determined by the power of the transmitter, the sensitivity of the receiver, and the transmission environment. Due to the UE transmission power being less than the eNB, the LTE network coverage is mainly determined by the UE transmission attenuation. From the 3GPP standard TS36.104 for LTE, the eNB reception radio sensitivity level is -93.5 dBm for the 5 MHz bandwidths [22]. We consider -93.5 dBm as the eNB radio sensitivity level and propagation loss model COST231 in ns-3 to simulate a suburban transmission environment. According to the configuration of UE power, eNB sensitivity level, and propagation model that we adopt, the transmission coverage for our experiment should be 327 m. We set all UEs within 300 m to rule out the impact of radio attenuation on simulation results. Our approach adopts 3 % PLR as the threshold. When the PLR for transmission is above 3 %, our approach will duplicate the data packets for TT nodes of this transmission. We executed the simulation five times and each execution time is 60 s. We adopt 95 % confidence intervals in our experiments.

B. Evaluation Results

Figs. 6, 7, and 8 show the performance of TT, ET, and BE data transmission (with respect to throughput, latency, and PLR) with increasing numbers of ET and BE devices in the TTP scenario. The TT nodes use the pre-defined dedicated time slots to transmit data. It ensures that TT traffic can have the highest priority on data transmission, followed by ET traffic, and then BE traffic. From Fig. 6, we observe that

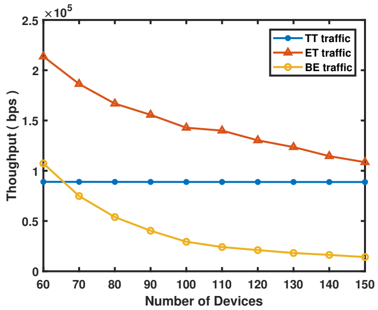

Fig. 6. Average Throughput

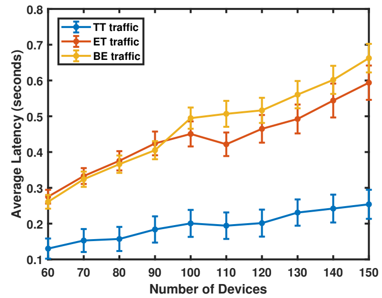

Fig. 7. Average Latency

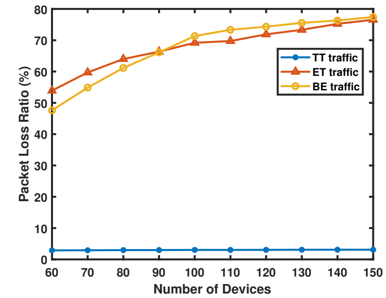

Fig. 8. Packet Loss Ratio

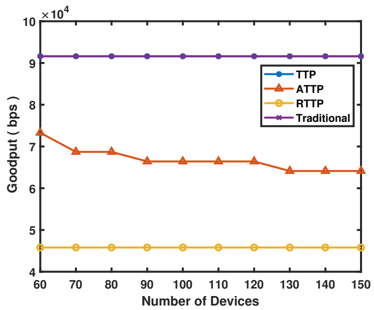

Fig. 9. Goodput

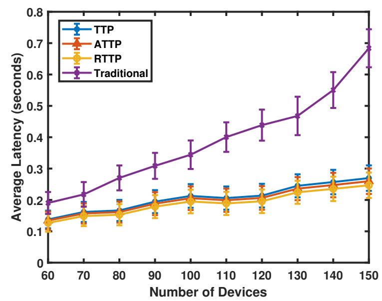

Fig. 10. Average Latency

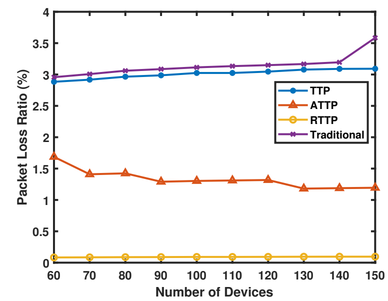

Fig. 11. Packet Loss Ratio

TABLE I
SIMULATION PARAMETERS

Parameter	Value
UE Tx Power	23 dBm
UE Height	1m
UE category	CAT-0
eNB Tx Power	46 dBm
eNB Height	30m
eNB Sensitivity Level	-93.5 dBm
UlEarfcn	18100
Uplink Central Band Frequency	1930.0 MHz
DlEarfcn	100
Downlink Central Band Frequency	2120.0 MHz
Antennas	SISO*
Uplink Data Rate	512 kbps
Transport Layer Protocol	UDP**
Uplink Bandwidth (MHz)	5
Number of Uplink Resource Block	25
Number of TT nodes	20
Simulation Time	60 s

* Single Input Single Output (SISO)
** User Datagram Protocol (UDP)

the throughput of TT traffic is stable even with the increased number of ET and BE nodes. Due to the TT nodes being assigned reserved time slots, and ET and BE nodes are not allowed to transmit their data within these time slots to avoid conflict with TT nodes. Fig. 7 shows the average transmission latency for the three types of data packets. The average latency of TT packets remains below 0.254 s. Even with the maximum standard deviation, latency is still less than 0.3 s. The average delay for ET and BE packets is much higher than TT traffic. Particularly, as BE traffic adopts the random back-off time to avoid conflicting with TT traffic and the bandwidth competition with ET traffic, the average delay of BE packets is the highest. Fig. 8 illustrates the PLR of the three types of traffic. We can see from the figure that the packet loss ratio of TT traffic is stable and remains low, between 2.884 % to 3.09 %. The packet loss ratios of ET and BE traffic increase rapidly along with the increase in number of IoT devices.

Fig. 9, 10, and Fig. 11 compare the transmission performance of TTP, RTTP, ATTP and Non-TTP. In TTP, exclusive time slots are reserved for all TT nodes to attain conflict-free real-time transmission. RTTP adopts the replication scheme with the fixed replication number for TT data packets to achieve reliability in data transmission. ATTP uses the dynamic replication scheme to dynamically adjust the repetition number for TT data packets according to channel quality, applying a higher number of repetitions for poorer quality connections, ensuring low PLR and real-time transmission. Fig. 10 shows that the average delay for Non-TTP grows rapidly along with the increase in number of IoT devices. In contrast, the average latency of the TT packets for the three TTP scenarios increased more slowly, especially for the scenarios where ATTP and RTTP are in place. This demonstrates that TTP obtains less transmission delay than the Non-TTP network. Among them, RTTP obtains the best performance (i.e., best real-time transmission).

Fig. 11 exhibits PLR for TT traffic on TTP, ATTP, RTTP, and Non-TTP. From the figure, we can see that the PLR of RTTP is less than 0.1 %. ATTP attains reasonable PLR at less than 1.7 %, and both TTP and Non-TTP observe PLR of approximately 3.5 % (Non-TTP slightly higher than TTP). From Fig. 9, the goodputs for TTP and Non-TTP are the same. Nonetheless, the goodput of RTTP is much smaller than TTP, as RTTP replicates all data packets. Such a repetition scheme results in only half of throughput carrying on useful information. Our ATTP, in contrast, simply replicates the data packets transmitted in poor communication channels, attaining higher goodputs by dynamic packet repetition. Our ATTP shows expected performance on real-time transmission with low data transmission latency and supports the transmission reliability via low PLR, with marginally reduced cost in goodputs.

Through our simulated results, we confirm that our ATTP and RTTP improve performance for TT traffic on packet transmission latency and PLR, at the cost of marginally lower goodputs, compared to TTP and Non-TTP. Moreover, the latencies of ATTP and RTTP are close, however, ATTP achieves higher goodput than RTTP. In our evaluation, the goodput using ATTP is 75 % of TTP, and the goodput using RTTP is 50 % of TTP. The evaluation results clearly show better performance for ATTP compared to RTTP in terms of overhead to the network.

VI. Final Remarks

In this paper, we proposed our ATTP approach that supports real-time data transmission in wireless networks. In particular, we designed a dynamic replication scheme to ensure the reliability and latency of Time-Triggered (TT) traffic delivery with only minimal additional overhead to the network. We also designed a multi-class traffic scheduling scheme that can efficiently allocate network resources for both TT traffic and non-TT traffic. Through extensive performance evaluation, our experimental results confirm the effectiveness of our approach with respect to latency, packet loss ratio, and overhead, compared with several baseline schemes. We found that a traditional non-TTP scheme has higher goodput, but higher latency and PLR. TTP has higher goodput also, comparable latency to ATTP, but high PLR. RTTP achieves low PLR but at the cost of low goodput. Thus ATTP gives the best overall performance.

References

[1] J. A. Stankovic. Research directions for the internet of things. *IEEE Internet of Things Journal*, 1(1):3–9, 2014.

[2] H. Xu, W. Yu, D. Griffith, and N. Golmie. A survey on industrial Internet of Things: A cyber-physical systems perspective. *IEEE Access*, 6:78238–78259, 2018.

[3] G. Xu, W. Yu, D. Griffith, N. Golmie, and P. Moulema. Toward integrating distributed energy resources and storage devices in smart grid. *IEEE Internet of Things Journal*, 4(1):192–204, Feb 2017.

[4] A. Zanella, N. Bui, A. Castellani, L. Vangelista, and M. Zorzi. Internet of things for smart cities. *IEEE Internet of Things Journal*, 1(1):22–32, Feb 2014.

[5] H. Kopetz, A. Ademaj, P. Grillinger, and K. Steinhammer. The time-triggered ethernet (TTE) design. In *Eighth IEEE International Symposium on Object-Oriented Real-Time Distributed Computing (ISORC)*, pages 22–33. IEEE, 2005.

[6] V. Mikolasek, A. Ademaj, and S. Racek. Segmentation of standard ethernet messages in the time-triggered ethernet. In *2008 IEEE International Conference on Emerging Technologies and Factory Automation*, pages 392–399, 2008.

[7] nsnam. ns-3 network simulator. https://www.nsnam.org/. Accessed November 4, 2018.

[8] *Time-Triggered Ethernet*, nov 2016.

[9] H. Kopetz and G. Grunsteidl. TTP-A time-triggered protocol for fault-tolerant real-time systems. In *Proc. of FTCS-23 The Twenty-Third International Symposium on Fault-Tolerant Computing*, pages 524–533. IEEE, 1993.

[10] T. Steinbach, F. Korf, and T. C. Schmidt. Comparing time-triggered Ethernet with FlexRay: An evaluation of competing approaches to real-time for in-vehicle networks. In *2010 IEEE International Workshop on Factory Communication Systems Proceedings*, pages 199–202. IEEE, 2010.

[11] C. Lu, A. Saifullah, B. Li, M. Sha, H. Gonzalez, D. Gunatilaka, C. Wu, L. Nie, and Y. Chen. Real-time wireless sensor-actuator networks for industrial cyber-physical systems. *Proceedings of the IEEE*, 104(5):1013–1024, 2016.

[12] W. Steiner. An evaluation of SMT-based schedule synthesis for time-triggered multi-hop networks. In *Proc. of 2010 31st IEEE Real-Time Systems Symposium*, pages 375–384. IEEE, 2010.

[13] F. Shan, W. Liang, J. Luo, and X. Shen. Network lifetime maximization for time-sensitive data gathering in wireless sensor networks. *Computer Networks*, 57(5):1063–1077, 2013.

[14] D. Shrestha, Z. Pang, and D. Dzung. Precise clock synchronization in high performance wireless communication for time sensitive networking. *IEEE Access*, 6:8944–8953, 2018.

[15] P. Bartolomeu, M. Alam, J. Ferreira, and J. Fonseca. Implementation and analysis of wireless flexible time-triggered protocol. *Ad Hoc Networks*, 58:36–53, 2017.

[16] Romain R. Jacob, L. Zhang, M. Zimmerling, J. Beutel, S. Chakraborty, and L. Thiele. TTW: A time-triggered wireless design for cps. In *Proc. of 2018 Design, Automation & Test in Europe Conference & Exhibition (DATE)*, pages 865–868. IEEE, 2018.

[17] H. Xu, X. Liu, W. Yu, D. Griffith, and N. Golmie. Reinforcement learning-based control and networking co-design for industrial internet of things. *IEEE Journal on Selected Areas in Communications*, 38(5):885–898, 2020.

[18] P. Bartolomeu and J. Fonseca. Towards flexible time triggered wireless communications. In *Proc. of 2010 IEEE International Workshop on Factory Communication Systems Proceedings*, pages 203–206. IEEE, 2010.

[19] J. Zhang, Y. Wu, W. Yu, and C. Lu. A QoS aware uplink scheduler for IoT in emergency over LTE/LTE-A networks. In *Proc. of International Conference on Software Engineering Research, Management and Applications*, pages 1–22. Springer, 2019.

[20] J. Zhang, W. Yu, and C. Lu. Performance assessment of LTE/LTE-A based wireless networks for internet-of-things. In *Proc. of 2019 IEEE/ACIS 18th International Conference on Computer and Information Science (ICIS)*, pages 18–23, 2019.

[21] sharetechnote. Communication channel model. https://www.sharetechnote.com/html/Communication_ChannelModel_Diversity_Rx.html/. Accessed May 9, 2020.

[22] Evolved universal terrestrial radio access (E-UTRA); base station (BS) radio transmission and reception. *3GPP Standard*, (TS36.104 release 12), 2015.