IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received March 21, 2019, accepted April 14, 2019, date of publication May 3, 2019, date of current version May 17, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2914658

Dynamic QoE/QoS-Aware Queuing for Heterogeneous Traffic in Smart Home

MAROUA BEN ATTIA, KIM-KHOA NGUYEN,
AND MOHAMED CHERIET, (Senior Member, IEEE)
École de Technologie Supérieure (ÉTS), Montreal, QC H3C 1K3, Canada

Corresponding author: Maroua Ben Attia (maroua.benattia@gmail.com)

This work was supported in part by the Natural Sciences and Engineering Research Council of Canada (NSERC), in part by the Ericsson under Project CRDPJ 469977, and in part by the Canada Research Chair, Tier 1.

ABSTRACT Smart home gateways have to forward multi-sourced network traffic generated with different distributions and with different quality-of-service (QoS) requirements. The state-of-the-art QoS-aware scheduling methods consider only the conventional priority metrics based on the IP type of service (ToS) field to make a decision for bandwidth allocation. Such priority-based scheduling methods are not optimal to provide both QoS and quality of experience (QoE), since higher priority traffic may not require lower delay than lower priority traffic (for example, traffic generated from medical sensors has a higher priority than traffic from streaming devices, but the latter one requires lower maximum delay). To solve the gaps between QoS and QoE, we propose a new queuing model for QoS-level Pair traffic with mixed arrival distributions in the smart home network (QP-SH) to make dynamic QoS-aware scheduling decisions meeting delay requirements of all traffic while preserving their degrees of criticality. A new metric that combines the ToS field and the maximum number of packets that can be processed by the system's service during the maximum required delay is defined. Our experiments show that the proposed solution increases 15% of packets that meet their priorities and 40% of packets that meet their maximum delays as well as 25% of the total number of packets in the system.

INDEX TERMS Quality of service, quality of experience, smart home, traffic scheduling optimization.

I. INTRODUCTION

A smart home network is a network that connects sensors, home appliances, and intelligent devices that react with each other with user instructions or system provider (for example remote control of devices or intelligent heating systems automatically adapting to outdoor temperature) [1]. Smart home networks are evolving rapidly to include heterogeneous physical access (both wired and wireless) and a large number of smart devices that generate different types of traffic with different distributions. Also, a variety of applications (VoIP, messaging, video, etc.) with different requirements is putting more constraints in smart home traffic scheduling such as congestion and delay. This requires automated management of traffic loads within the home gateway by offering more than one priority class. From the perspective of Internet Service Providers (ISP), these classes are defined based on bandwidth requirements for critical applications using IP ToS field [2]. However, from home user's perspective the priority classes correspond to the delay of traffic, especially for video streaming applications. For example, packets generated from a fire detector or medical sensors have a higher priority than packets generated from streaming devices. On the other hand, streaming devices with video bitrates from 400 kbps to 14,000 kbps [3] require a lower maximum delay compared to periodic sensing objects such as medical sensors (with sensing rate between 12 bps and 12 kbps [4]). Thus, traffic scheduling in smart home network should consider specific Quality of Experience (QoE) metrics of each type of traffic in addition to the conventional Quality of Service (QoS) metric which is based on IP ToS field or user preferences. Traffic from each application must be mapped to both priority class and delay-sensitive class and processed by a proper scheduling discipline to meet both criticality and delay requirements to avoid local network congestion. The most challenging issue faced by smart home gateway is to

The associate editor coordinating the review of this manuscript and approving it for publication was Monjur Mourshed.

provide both ISP and home users satisfactions in terms of QoS and QoE especially for delay-sensitive applications [5]–[8] through finding an automatic way to schedule multi-sourced packets while considering their degree of criticality and meeting their maximum required delay. Most of the prior work on QoS scheduling problem [9]–[19] cannot be efficiently applied in smart home network since they do not consider the impact that prioritizing specific traffic based only on static metrics like TOS field or user-defined preferences may have on other network traffic (lower-priority traffic may miss their maximum allowed delay when prioritizing higher priority traffic having a higher upper-delay bound).

In this paper, we propose a dynamic model for optimizing packet scheduling in the smart home network with mixed arrival distributions while considering both the critical nature of application traffic and its maximum allowed delay. The contribution of this paper includes a new dynamic queuing model for smart home network traffic generated by heterogeneous sources, which increases the number of processed packets that meet their deadline and preserves their degree of criticality. The rest of the paper is organized as follows. Related studies on QoS based scheduling is discussed in Section II. In Section III, we will describe the smart home traffic scheduling with QoS constraints. QoS scheduling problem is presented in Section IV. Section V describes the proposed queuing model for QoS-level Pair Heterogeneous-sourced traffic in the smart home network (QP-SH). Numerical results of our solution are provided in Section VI. Finally, we draw conclusions and present future work.

II. RELATED WORK

Many scheduling algorithms have been proposed in previous work to manage different type of network traffic (resumed in Table.1). Benacer *et al.* [9] contributed a high capacity hybrid Priority Queuing (HPQ) for high-speed network devices. HPQ is a fixed priority algorithm based on Priority Queuing (PQ), which considers the priority order of inserting packets. Shakir and Rajesh [10] contributed a two-level queuing model that considers the theoretical delay to provide QoS requirements in LTE networks. In the first layer queuing, packets are sorted based on their size, their expected departure time and the service time; then, they are scheduled to forme calendar discs using a weighted fair queuing algorithm (WFQ). In the second layer queuing, the calendar discs are sorted based on their frequency bands and their corresponding packets are selected using Weighted round-robin algorithm (WRR) a generated form of Fair Queuing (FQ), which allows, at each scheduling round, en/dequeuing a certain number of packets (weights) from each queue. Anand and de Veciana [11] contributed a multi-class scheduler which optimizes end-user QoE based on mean flow delay in wireless networks. Their solution uses a weighted Gittins index scheduler to optimize resources allocation for different classes of applications according to their sensitivity towards the mean delay. Bakhshi and Ghita [12] proposed a queuing model that considers user-defined profile priorities to optimize bandwidth allocation in-home network. Their solution is based on Software Defined Network (SDN) technology to calculate user-profiles in a central controller which resides on the cloud and push the resulting rules on a home gateway. The authors evaluate their solution using multimedia and video streaming applications. Their solution has shown good performance in terms of latency and packet loss for only a selected set of high priority users.

TABLE 1. Related work.

ref	QoS	QoE	Applications/Scope
[9]	The order of inserting packets	None	Wireless networks
[10]	Delay, service time and packet size	None	LTE networks
[11]	None	Mean flow delay	Wireless networks
[18]	The global throughput	Delay	Wireless networks
[20]	None	The duration of video playback interruption	Video-on-demand applications in Wireless network
[12]	Bandwidth	User-defined profiles	Multimedia and video streaming applications
[21]	the average end-to-end delay	The frame buffer level at the destination nodes	Video Streaming
[14]	Packet inspection	None	Video Streaming
[15]	Real-time bandwidth	None	Video Streaming
[16]	None	Energy consumption	Vehicular networks
[17]	Energy consumption	Loss tolerance	Loss tolerant App. in 5G
[19]	None	Bandwidth allocation	Home network
[16]	Current active applica-tion/device	None	Home network

Bozkurt and Benson [13] contributed a context-aware scheduling discipline which prioritizes home network traffics based on the currently active applications and devices. Yang *et al.* [14] proposed a cloud-based scheduling solution to prioritize home applications using packet inspection. The authors evaluate their solution using video streaming applications. Their architecture risk to let queues of the low-high priorities starve since it's considers only the static nature of priority assignments. Abuteir *et al.* [15] contributed a Wireless Network Assisted Video Streaming (WNAVS) framework which relies on SDN technology to schedule home packets based on real-time bandwidth allocation and network traffic statistics. However, their solution focuses only on one type of home application which is not the case for real home network traffic. Hsieh and Hou [20] proposed an online schedule which maximizes wireless network utility based on the QoE of each flow. The authors used the duration of video playback interruption to optimize QoE for video-on-demand applications under heavy-traffic conditions. Their solution proposed to schedule the client with the largest data rate in each scheduling period if there are no ties. If a tier occurs, the selected client is the one with the smallest product of its weight and the difference between the total amount of received data and the total number of bits that should have been played if there is no video interruption. Each client is assigned a weight by the access point that reflects its class of service.

Zeng *et al.* [16] contributed a scheduling scheme for Vehicle Ad hoc NETworks which increases the QoE of charging and discharging electric vehicles while optimizing the load capacity of the power grid. Each electric vehicle is matched to the charging station that maximizes its charging utility and has at least one free interface. Electric vehicles may cooperate in the same charging station by selling their electricities (discharging) to vehicles with low battery levels. The cooperative electric vehicles charging and discharging is scheduled using a Pareto Optimal Matching Algorithm. Butt *et al.* [17] proposed a cross-layer scheduling framework over fading channels which guarantees the minimum QoS requirements in terms of energy consumption while satisfying the QoE in terms of loss tolerance for loss-tolerant applications in the 5G wireless network. The authors used the Markov decision process to model their scheduling problem, and they used stochastic optimization techniques to solve it. Zheng *et al.* [18] proposed a task layer scheduling scheme to improve QoE in terms of the quality of the transmission of a group of packets (called task) rather than the quality of the link in wireless networks. Each link can support many tasks from a different class of services with different delay constraints. Their solution calculates the remaining time of each task and each link. Then, the link with the least remaining time is selected to schedule tasks with the fewest packets. Authors considered the QoE using the global throughput and the QoS using maximum delay for each class of service. Fan and Zhao [21] contributed a cross-layer scheduling scheme for video streaming which considers the average end-to-end delay and the frame buffer level at the destination nodes to improve both QoS and QoE in wireless Ad hoc networks. The authors used the Lyapunov optimization framework to solve the optimization problem and proposed a distributed media access control algorithm to reduce computational complexity.

Chaabnia and Meddeb [19] contributed a new distributed model for home network traffic prioritization based on SDN technology. The authors implemented two-level slicing strategies; control-level slicing where traffic is prioritized based on bandwidth requirements and data plane level slicing where traffic is prioritized based on the type of application. Each data plane slice is associated with one control plane slice. The authors evaluate three scenarios of their solution; same priority slices, ascending order priority slices and descending order priority slices (referring to PQ). Packets with low priority in the second and third scenarios may suffer from the starvation problem.

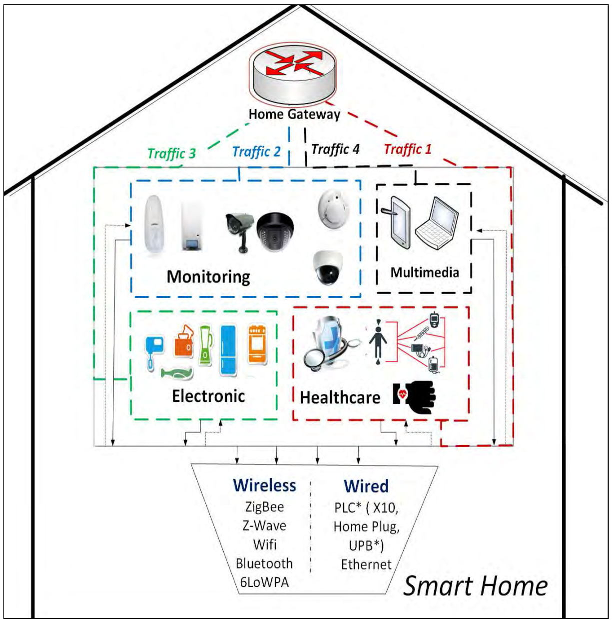


FIGURE 1. **Smart home network.**

In general, most of the existing scheduling solutions rely on static metrics in the priority assignment task. They are either based on user-defined profiles, current active applications or class of service. Even though there are solutions that assign priorities dynamically (based on real-time bandwidth allocation or source-destination distance), they consider a specific type of home application (multimedia and video streaming applications) or only a particular optimization goal. They either focus on improving QoS from the perspective of ISP (optimize bandwidth utilization based on traffic loads to meet ToS priorities) or improving QoE from the perspective of the home user (optimize delay based on the distance between the source and destination nodes).

Specific queuing metrics, which need to be determined in the smart home network, like traffic application criticality (or type of service) and the maximum required delay along with heterogeneous distributions queuing adaptability, has never been taken into account. These factors are very important in the context of the home network to fill the gap between QoS and QoE for any home application in an automated way. Our approach mitigates these limitations by considering these important key factors to deploy a new scheduling scheme specific to the smart-home network context. More specifically:

- Proposing a new deterministic queuing model for multi-sourced traffic generated with different distributions using a new composite QoS-level metric based on both criticality-based priority and delay-based priority to avoid local network congestion by optimizing the number of packets that meet their allowed delay while preserving their degree of criticality.

III. SYSTEM DESCRIPTION

Fig. 1 depicts a typical smart home network. Each home network includes many different multimedia devices (i.e., tablets, smart-phones, connected TVs, etc.) and objects

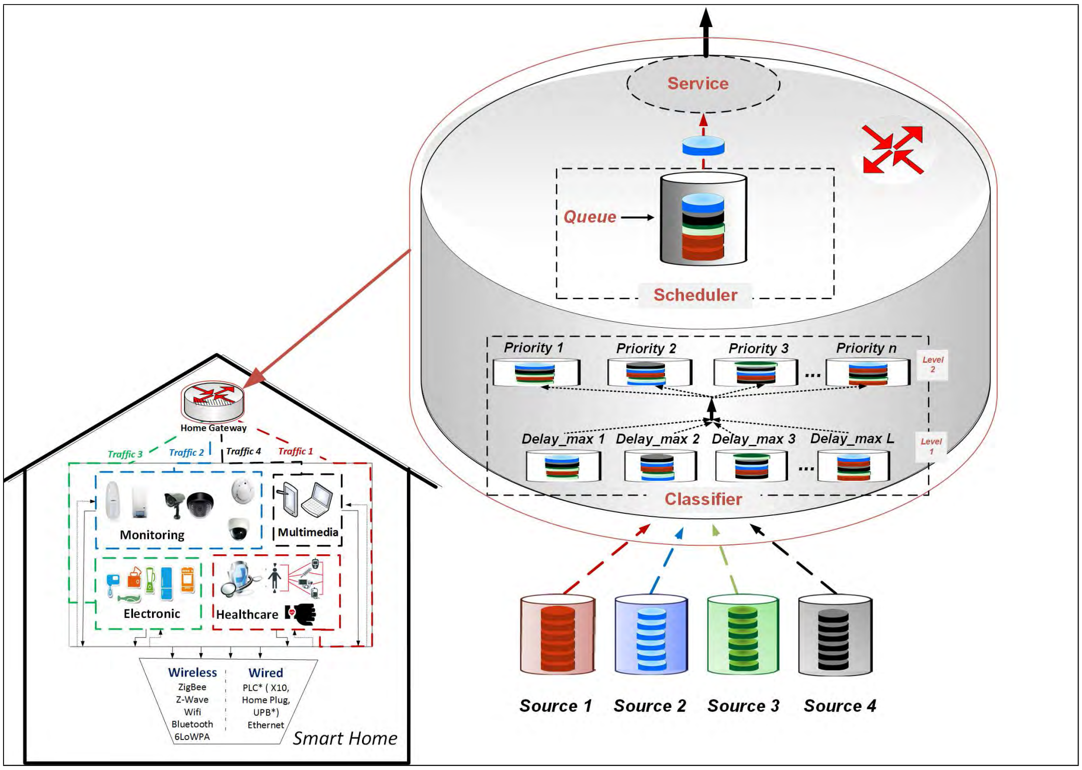


FIGURE 2. System description.

(i.e., sensors, electronics, appliances, etc.). Sensors are devices used to detect the location of people and objects or to collect data or states (i.e., temperature, energy consumption, open windows/doors, movement, broken glass). Electronic devices include phones, televisions, and laptops. Electrical devices refer to toasters, kettles, light bulbs, etc. Appliances refer to washing machines, refrigerators, etc. Such a network offers services to a wide range of application like monitoring, health assistance, safety, and energy efficiency, producing traffic with different Quality of Service (QoS) levels [22]–[24](marked by different colors in Fig. 1) and managed by the smart home gateway.

Fig. 2 illustrates an example of the smart home gateway. The home gateway contains three modules [25]; Classifier, Scheduler, and Service. in this paper, we use two-level classifier which classifies the network packets firstly according to their maximum allowed delay and then, according to their priorities. Scheduler contains the queue in which classified packets will be scheduled according to their arrived time and their two-level priorities. The number of priority classes n is fixed. It is calculated based on both the heterogeneity of constraints imposed by the traffic data and the maximum available bandwidth in the system. A small value of n may increase the available bandwidth while fulfilling fewer constraints with a partial QoS hierarchy. However, a high value of n may increase bandwidth utilization while satisfying QoS requirements for a large number of data type. Service module contains c parallel servers. We assume that the main queue of the system has an unlimited size (storage area) as long as the service module can process up to c packets per service time using its parallel servers. In fact, given the limited number of smart home applications, where each home may have 10 devices, and the processing capacity of the smart home gateway, which can process up to 100 000 flows in the network [26], we can assume that the main queue has unlimited size.

A. IMPLEMENTATION MODEL

Smart home network enables multiple smart objects to operate in one home gateway. Each network flow is assigned a priority group to prioritize their traffic by QoS packet marking using ToS (or DS) bits in the IP header [27]. On the other hand, each application is assigned a maximum allowed delay D_{max} that has to be met by their packets. Home gateway schedules network traffics firstly using w_{max} metric (see section IV-B) calculated based on their maximum delay D_{max} and then, using ToS field based on their assigned QoS priorities to provide both QoE and QoS in the smart home network. In our proposed architecture, a simple modification on the IP protocol stack will be made by encapsulating a new field in the IP header that reflects the maximum allowed delay D_{max} for each packet besides ToS field.

The problem we address in this paper is to provide optimal scheduling for packets generated from different sources and with varying distributions with respect to their delay budget and their degree of criticality.

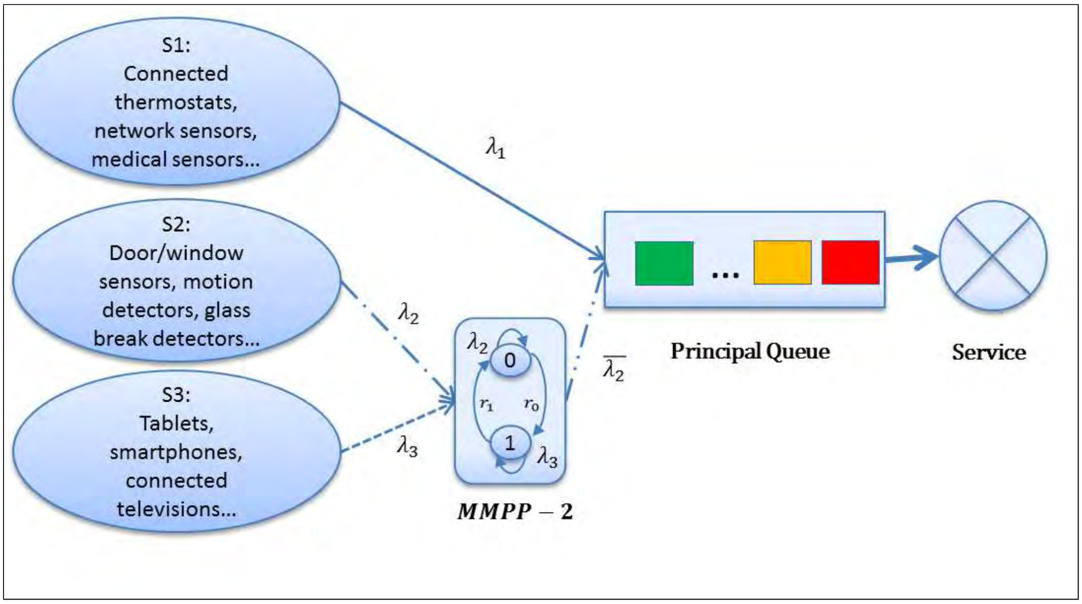


FIGURE 3. Modeling the input traffic.

IV. QOS-AWARE SCHEDULING PROBLEM

Our problem is optimizing QoS scheduling for smart home network traffic. It consists of finding a way to schedule multi-sourced packets, that ensures their maximum tolerated delay and preserves their degree of criticality. The contribution of this paper is improving previous work by introducing a dynamic QoS level pair for multi-sourced traffic with different arrival rate, that considers the criticality of the application all along with the maximum number of packets that can be processed before processing the packet based on its maximum tolerated delay.

A. MODELING AND CHARACTERIZING THE INPUT TRAFFIC AND THE SERVICE

Incoming traffic can follow different distributions depending on their data type as well as the type of their generation process (or source Si) as described in Fig. 3:

a: PERIODIC SENSING OBJECTS (S_1)

These objects periodically detect and send to a central server (usually on the cloud) the states of monitored devices for each period T (i.e., connected thermostats, network sensors, medical sensors, etc.). A packet should be sent by sensors every period T and sent out by the gateway before 2T (the time when the following packet arrives). This type of source generates discrete traffic, with each period T (synchronous) and with a constant, determined distribution (D).

b: EVENT-TRIGGERED SENSING OBJECTS (S_2)

These objects generate traffic by triggering some events (for example, door/window sensors, motion detectors, etc.) to indicate the status of the monitored object or person. Sensing data are delay-sensitive tasks that must be processed quickly to prevent serious property damage or injury since a small fire can rapidly turn fatal and we not always have enough time for safe evacuation. We define $D^{q_i}_{max}$ the maximum tolerated delay for QoS-level q_i traffic. The generation of this traffic is generally rare and does not depend on any other traffic (decorrelated). The arrival of this type of traffic (average arrival number λ_2)) can, therefore, be modeled according to a distribution of the Poisson process with an exponential inter-arrival rate (M).

c: STREAMING OBJECTS (S_3)

These objects generate a continuous data stream (by tablets, connected televisions, surveillance cameras, etc.). These data do not always require QoS, however, for delay-sensitive applications like VOIP and video streaming (security camera or films), data should not be delayed to provide QoE (Quality of Experience) or security to the end user. Thus, the maximum tolerated delay for QoS-level q_i traffic generated from these type of objects is $D^{q_i}_{max}$. For video streaming applications, the maximum tolerated delay may increase as the frame rate decreases. Thus, the value of $D^{q_i}_{max}$ depends on the application requirements. For example, in video surveillance systems 7.5 frames per second (fps) are enough to capture and pause specific frames without noticing loss with the human eye [28], [29]. However, next-generation video devices like ultra-high definition TV (UHD), in which motion are often present, require a higher frame rate with a minimum of 60 fps [30]. Thus, the minimum required frame rate depends on the contents of the video. The higher the frames, the smoother the video will be. The generated data may reach peaks during periods of heavy use or may be negligible (like traffic from surveillance cameras or during the rest of the day). We have modeled this type of traffic with a binary Markov-Modulated Poisson Process (MMPP):

- State 0: incoming traffic follows a Poisson process with a very high arrival rate λ_3 ($\lambda_3 \gg \lambda_2$). This traffic corresponds to the flows generated during peak periods of use.
- State 1: incoming traffic follows a Poisson process with a low arrival rate λ_{31} ($\lambda_{31} \ll \lambda_3$). This traffic corresponds to the negligible flows generated during the rest of the day or by surveillance cameras.

The packet rate λ_2 generated by the source S_2 and the packet rate λ_{31} of the state 1 of the source S_3 are generally similar, and they can, therefore, be modeled by the same distribution with the same arrival rate λ_2. So in the following model, we replace all λ_{31} by λ_2.

The arrival flow of our system therefore follows two different distributions; a predetermined distribution with an arrival rate λ_1 and a binary Markov distribution with an arrival rate ($\overline{\lambda_2}$). If we consider $Pr(s = i)$ the probability that an arrival packet is in state i (with $i \in \{0, 1\}$) then we have:

$$\overline{\lambda_2} = Pr(s = 0)\lambda_2 + Pr(s = 1)\lambda_3 \tag{1}$$

$$Pr(s = 0) = \frac{r_1}{(r_0 + r_1)} \tag{2}$$

$$Pr(s = 1) = \frac{r_0}{(r_0 + r_1)} \tag{3}$$

with r_0 and r_1 are respectively the average lengths of stay in the state 0 and state 1 and therefore the arrival rate will be

$$\overline{\lambda_2} = \frac{\lambda_2 * r_1 + \lambda_3 * r_0}{(r_0 + r_1)} \tag{4}$$

We have a single domestic gateway with c servers. A server can process any packet with a size up to the Maximum Transmission Unit (MTU). We assume that all packets are MTU-sized packets and the service follows a deterministic distribution with a rate $\frac{1}{s}$.

B. MODELING QOS REQUIREMENTS FOR SMART HOME NETWORK DEVICES

For each smart home network application, we define a QoS level based on two main QoS parameters; a priority level and a maximum required delay. Priority level depends on the degree of the application criticality. Exceeding delay for critical applications is fatal, however, for non-critical applications, it is better to meet the deadline, but it is no crucial. For example, the processing time of packets generated from a fire detector must not exceed their maximum required delay otherwise the fire will rapidly turn fatal, however, a high processing time of a packet from video streaming applications, that exceeds its required maximum delay, will deteriorate the service without causing a real disaster. In our proposed architecture, three primary sources of traffic are considered (as described in section IV-A and as shown in Fig. 3); type 1 sensor S_1, type 2 sensors S_2 and multimedia devices S_3, along with only one home gateway. Each source can generate different QoS-levels of network traffic at different time slots, and a maximum of c packets can be processed at each service time s using c parallel servers having the same capacity (each server can serve up to one packet in s time slots).

This model is motivated as follows:

- General distribution function G for the service time,
- c servers: parallel servers having the same capacity,
- $D/G/c$ for traffic generated from source of type S_1: the interarrival time of data generated from periodic sensing objects S_1 is equal to a constant period of time and then, deterministic D (IV-A0a),
- $MMPP - 2/G/c$ for traffic from sources of type S_2 and S_3: the data generated from S_2 and S_3 are modeled with a binary Markov-Modulated Poisson Process (IV-A0c).

The service can serve:

- Up to c packets in s time slots,
- Up to $\frac{c}{s}$ packets in one time slot,
- Up to $\frac{c*D_{max}(P_i)}{s}$ packets during the maximum required delay $D_{max}(P_i)$ of a packet P_i.

Thus, for each packet P_i we define a maximum window size $w_{max}^{P_i}$ as the maximum number of packets that can be processed by the system's service during its required delay $D_{max}(P_i)$ as follows:

$$w_{max}^{P_i} = \frac{c * D_{max}(P_i)}{s} \tag{5}$$

We define the QoS-level pair q^{P_i}, for each network packet P_i as follows:

$$q^{P_i} = (p^{P_i}, w_{max}^{P_i}) \tag{6}$$

With p^{P_i} is the priority level of the P_i's application type.

As described in Fig. 4, we set a queue F^{q^i} for each QoS-level pair q^i and a scheduling discipline $D_F(F^{q^i})$ for composite QoS level packets from different F^{q^i} queues that we will determine later. We define a delay function for each packet $P^{(q,g)}$ generated from source S_i and having the QoS-level pair q as follow:

$$D_T(P^{(q,S_i)}) = \alpha_T(P^{(q,S_i)}) + s \tag{7}$$

With $\alpha_T(P^{(q,S_i)})$ is the waiting time of the packet $P^{(q,S_i)}$ before being served and s is the service time.

All used parameters and functions are listed in Table.2.

The smart home network is a heterogeneous infrastructure made of multiple electronic and electrical network devices like sensors, detectors, and laptops. These data sources generate a wide range of traffic with different distributions and various QoS and QoE requirements. A key challenge of this problem is to find a reasonable way to schedule multi-sourced packets from a composite class of service with respect to their QoS and QoE requirements. Thus, to meet the delay constraint, the delay of a packet $P_{ij}^{(q,S)}$ must be lower than the delay budget D_{max}^q required by the pair of the class of service q:

$$D_T(P_{ij}^{(q,S)}) \leq D_{max}^q \tag{8}$$

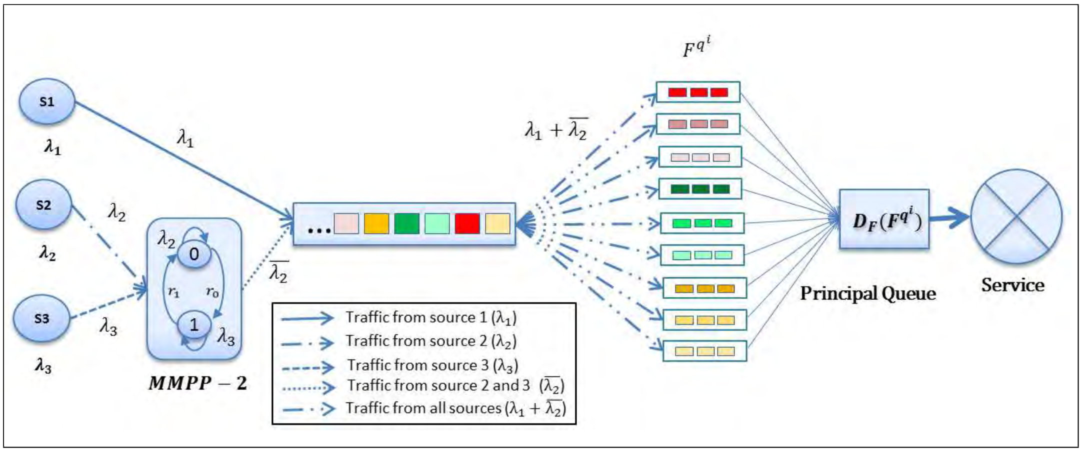


FIGURE 4. Composite QoS-level scheduling model.

TABLE 2. Notations.

Notations	Definitions
$S = \{S_i, i = 1, 2, \}$	Set of source of traffic in smart home
$F = \{F_i, i = 1, 2, \}$	Set of queues in the system
$Q = q^i$	Set of QoS-level pair q^i
$q^{(P_i)} = (p^{(P_i)}, w_{max}^{(P_i)})$	QoS-level pair of network packet P_i
$p^{(P_i)}$	Priority level of P_i's application type
$w_{max}^{(P_i)}$	Maximum number of packets that can be processed by the system's service before processing P_i
$D_{max}(P_i)$	Maximum required delay of P_i
$P = \{P_i^{(q,S_k)}\}$	Set of flows of QoS-level pair q and generated by source S_k
$P_i^{(q,S_k)} = \{P_{ij}^{(q,S_k)}\}$, $P_{ij}^{(q,S_k)}$	Flow i (of QoS-level pair q and from source S_k) and packet j of flow i
$D_F(F^{(q^i)})$	Scheduling discipline for composite QoS level queues
$\alpha_T(P_{ij}^{(q,S_k)})$	Waiting time function of packet $P_{ij}^{(q,S_k)}$ in the system
$D_T(P_{ij}^{(qS_k)})$	Delay function of packet $P_{ij}^{(q,S_k)}$ in the system

The QoS-aware scheduling problem consists of finding an optimal way to schedule packets from multi-sourced traffic with dynamic QoS-level pair that ensures the maximum tolerated delay and preserves their degree of criticality. We formulate the QoS-aware scheduling problem by the following objective function:

$$(D_F(F^{q^i}))^* = argmin_{D_T(P_{ij}^{(q,g)})} \begin{cases} P_{ij}^{(q,g)} \in P \\ D_T(P_{ij}^{(q,g)}) \leq D_{max}^q \end{cases} \tag{9}$$

V. QP-SH: QUEUING MODEL FOR QOS-LEVEL PAIR TRAFFIC IN SMART HOME NETWORK

To solve the queuing problem of smart home traffic that have a composite class of service $q^i = (p^i, w_{max}^i)$ and generated with different distributions, we propose a QP scheduling model as described in Fig. 5. The QP model dedicates a QoS-level pair $q^i = (p^i, w_{max}^i)$ for each packet generated from the different source of traffic. All packets with the same w_{max} will be merged to a single queue with the same w_{max} until reaching the main queue of the system. Then, packets in the same w_{max} queue will be scheduled according to their priority level p to ensure that each packet is processed according to its QoS-level pair whatever its source. In our proposed architecture, three main traffic sources are considered, as described in section IV-A and as shown in Fig. 4; type 1 sensors (S_1), type 2 sensors (S_2) and multimedia devices (S_3). Each source S_i has a set $F^W(S_j)$ of L queues for each w_{max}^i traffic generated from it with the rate λ_j, $F^W(S_j) = F^{(w_{max})}(S_j)$, $w_{max} \in W \subset F$ with W is the set of w_{max}.

Traffic from S_2 and S_3 are then modeled by a binary MMPP while keeping their priorities queues. All same w_{max} queues are merged to a single queue with the arrival rate ($\overline{\lambda_2}$). Then, all the same w_{max} queues from MMPP and S_1 are merged again to a single queue and sending to the principal queue with the arrival rate $\lambda_1 + \overline{\lambda_2}$ and $F^q = F^{(p,w)}$, $w \in W, p \in Q \in F^W \subset F$.

The QP-SH scheduling discipline is illustrated in Algorithm 1. The algorithm first initializes its queue using *init* function (Algorithm 2). Each arriving packet $P_{ij}^{((p^k, w_{max}^i), S_k)}$ generated from source S_k, is mapped to the queue $F^{w_{max}^i}(S_k)$ dedicated to its source. Packets with the same arrival time are pushed randomly at the queue $F^{w_{max}^i}(S_k)$. Then, all $F^{w_{max}^i}(S_k)$ queues from different sources of traffic will be merged to a single $F^{w_{max}^i}$ queue, packet per packet, based on their arriving times. Packets with the same

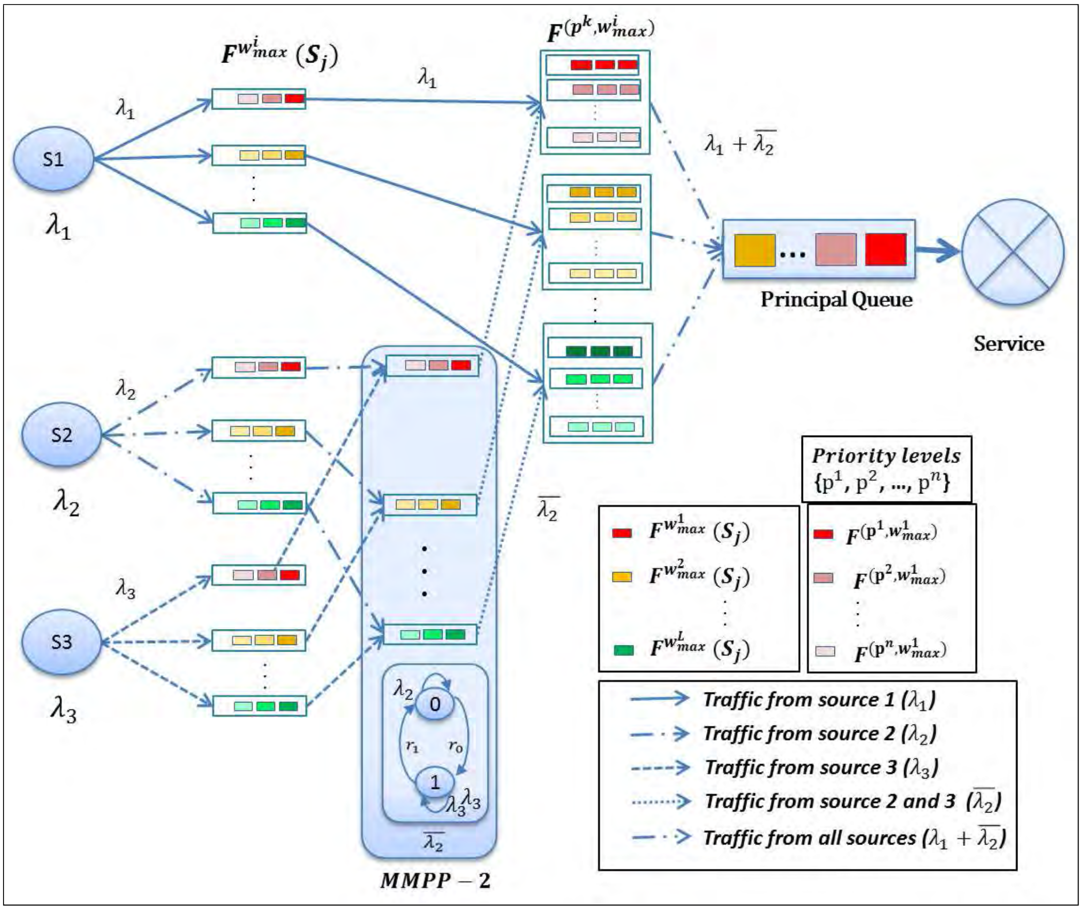


FIGURE 5. QP scheduling model.

Algorithm 1 QP-SH

```
procedure QP-SH($P$, $F$)
    init $P$, $F$
    $k = c$ // number of servers
    while $F^{w^*_{max}} = min_{w^i_{max}}(F^{w^i_{max}})$ non empty and $k! = 0$ do
        while $F^{p^{l^*},w^*_{max}} = min_l(F^{p^l,w^*_{max}})$, $l \in [1, n]$ non empty and $k! = 0$ do
            pull(FIFO($F^{p^{l^*},w^*_{max}}$))
            $k = k - 1$
        end while
        update($F$, $k$)
    end while
end procedure
```

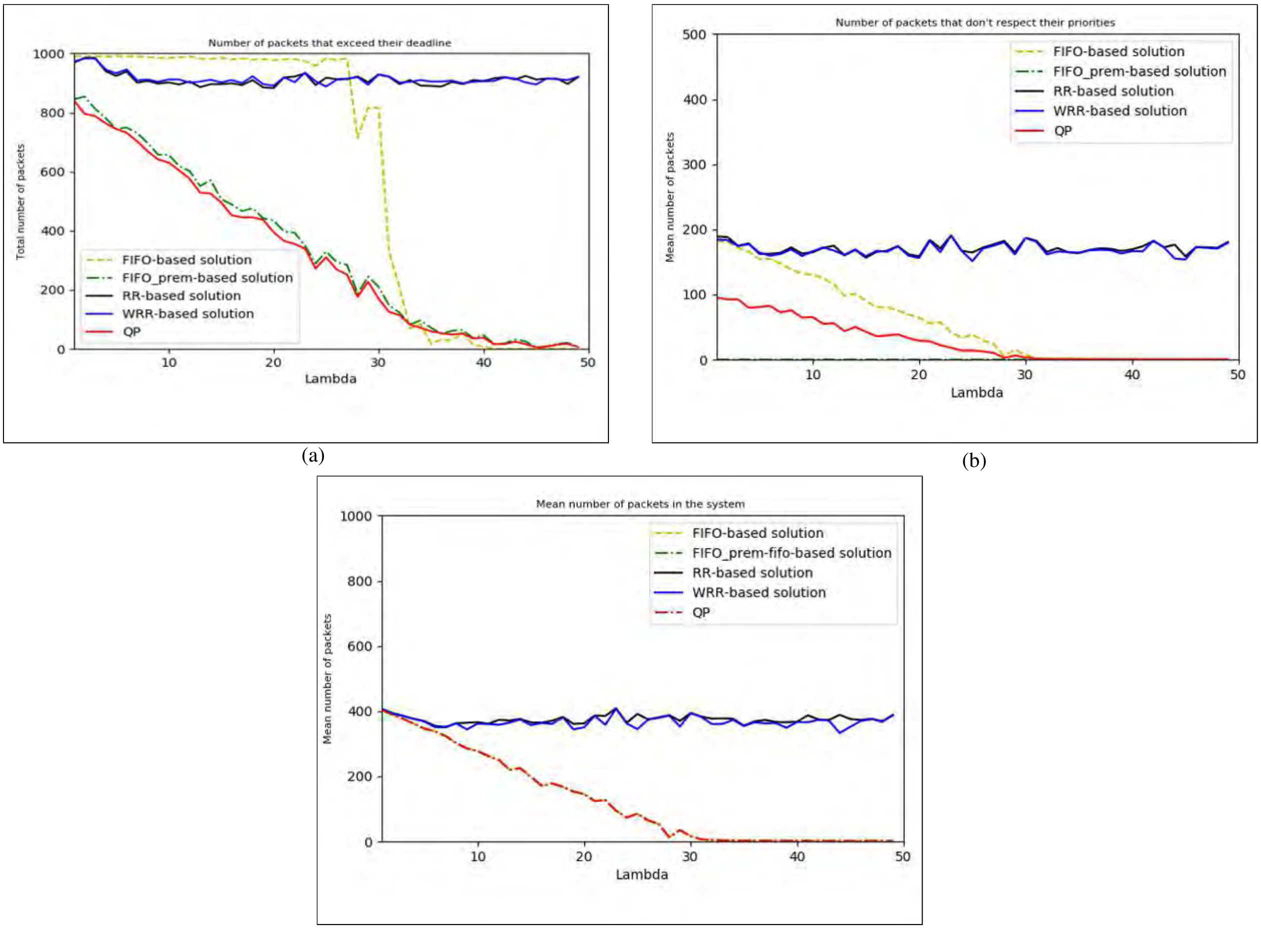
w^i_{max}, from different sources and having the same arrival time are pushed randomely to their corresponding queues $F^{w^i_{max}}$. All $F^{w^i_{max}}$ queues form a set $F = F^{w^i_{max}}, i \in L$ of queues. Then, all packets in each $F^{w_{max}{}^i}$ queue are grouped by priority into n sub-queues $F^{(p^k,w^i_{max})}$. The system processes all $F^{w^i_{max}}$ queuing in an ascending order beginning from the group of queue with the lowest w^i_{max} (Algorithm 1). Packets within the samew^i_{max} group are scheduling according to their priorities; packets highest priority are served first.

After each service round, the value of w^i_{max} for all $i \in L$ is decremented by the number of served packets (up to c packets since we have c servers), as the number of packets that can be processed by the system's service before processing each packet $P_{ij}{}^{((p^k,w^i_{max}),S_k)}$ is decremented by the number of served packets (see Algorithm 3).

VI. PERFORMANCE EVALUATION

To evaluate the performance of the proposed QP-SH algorithm, we build a simulation with up to 1000 network packets generated with different distributions and a different number of servers ($c = 1 - 5$). The D/G/c model is simulated using traffic generated (from periodic sensing objects) each 5ms with a rate 1/5 packet/ms ($\lambda_1 = 1/5$). Incoming traffic from event-triggered sensing objects follow exponential distribution with a rate $\lambda_2 = 0.5 * \lambda_3$ since it is much lower than λ_3 (as described in section IV-A). This negligible traffic is generated during $r_0 = 40\%$ of the day.

Incoming traffic from streaming objects follows an exponential distribution with a rate of λ_3 set from 1 to 50 packet/ms. This traffic is generated during periods of



(a)

(b)

(c)

FIGURE 6. **QP-SH performances in function of the arrival rate *lambda* (the service time is fixed to 30 ms). (a) Percentage of packets that exceed their maximum delays (%). (b) Percentage of packets that do not respect their priorities (%). (c) Mean number of packets in the system.**

heavy use, during $r_1 = 20\%$ of the day. We calculate $\overline{\lambda_2}$ as defined in Eq.(1). We randomly set the packet priority and the maximum delay. Regarding the service time, we consider two scenarios; in the first scenario, the service can serve a packet in 30 ms with a rate of 2 packet/s, and, in the second scenario, the service time varies from 10 ms to 60 ms. In both scenarios, we calculate the performance parameters of the global queue scheduling model based on the arrival rate $lambda = \overline{\lambda_2} + \lambda_1$ (where $\overline{\lambda_2} = \frac{r_0}{(r_0+r_1)}\lambda_2 + \frac{r_1}{(r_0+r_1)}\lambda_3$ as defined in Eq.(1), Eq.(2) and Eq.(3)). The different values of *lambda* are obtained by varying λ_3 from 1 to 50 and λ_2 in function of λ_3 ($\lambda_2 = 0.5 * \lambda_3$). Table.3 describes our experimental setup.

In Fig. 6, we consider the first scenario where the service time is fixed and we plot the curves of the number of packets that exceed their maximum delays (Fig.6(a)), the mean number of packets that do not respect their priorities (Fig.6(b)), and the mean number of packets in the system (Fig.6(c)) in function of the arrival rate *lambda*. These results are obtained using our QP-SH algorithm, the existing Round-Robin(RR) [31] and Weighted RR (WRR) [32] based solutions and the existing First in First out (FIFO) and FIFO preemptive (FIFO-prem) based solutions [9]. The mean number of packets that do not respect their priorities can be identified by comparing the QoS-level pair classification method (which is based on the priority provided by the two-level classifier; first using the maximum allowed delay and then, using QoS priorities) with that based on the QoS priority provided by the ToS field in the IP header.

TABLE 3. **Experimental setup.**

Parameter	Value
Number of packets	1000
D_{max}	uniform(200,250) (ms)
Priority	randint(0,10)
λ_1	1/5 (packet/ms)
λ_2	0.5 * λ_3 (packet/ms)
λ_3	1-50 (packet/ms)
Service time	30 ms (in scenario 1), 10-60 ms (in scenario 2)
Number of servers c	1-5
r_0	40 (%)
r_1	20 (%)

We note that the curves obtained with QP-SH algorithm are under the curves obtained with the RR, WRR, FIFO, and FIFO preemptive based solutions for the majority of criteria. We also note that the number of QP-SH based packets that violate their maximum delay and do not respect priority criterion decreases when we increase the arrival rate up to

Algorithm 2 Init

```
function INIT($P, F$)
    $//F = \{F^{w^i_{max}}, i \in L\}$
    $//F^{w^i_{max}} = \{F^{(p^k, w^i_{max})}(S_j), 0 \le k \le n, S_j \in S\}$
    $//P = \{P_{ij}^{(q,S_k)}\}$
    while arriving packets at time slot $t = P^t$ do
        for each $P_{ij}^{(q,S_k)} inP^t$ do
            push($P^{((p^k,w^i_{max}),S_k)}$, $F^{w^i_{max}}(S_k)$)
        end for
        $F = \cup_{i \in L} F^{w^i_{max}}$
    end while
    for each $P_{ij}^{(p^k,w^i_{max})} \in F^{w^i_{max}}$ do
        push ($P_{ij}^{(p^k,w^i_{max})}$, $F^{(p^k,w^i_{max})}$)
        $//F^{(p^k,w^i_{max})} = F^{q^{ij}}$
    end for
end function
```

Algorithm 3 Update

```
function UPDATE($F, k$)
    for each $F^{w^i_{max}} inF$ do
        $F^{w^i_{max}} = F^{w^i_{max} - k}$
    end for
end function
```

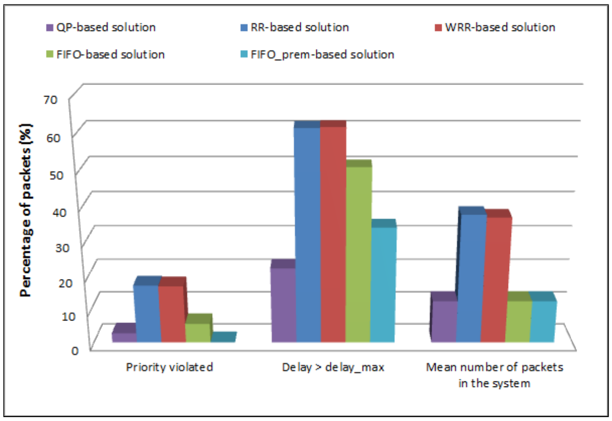
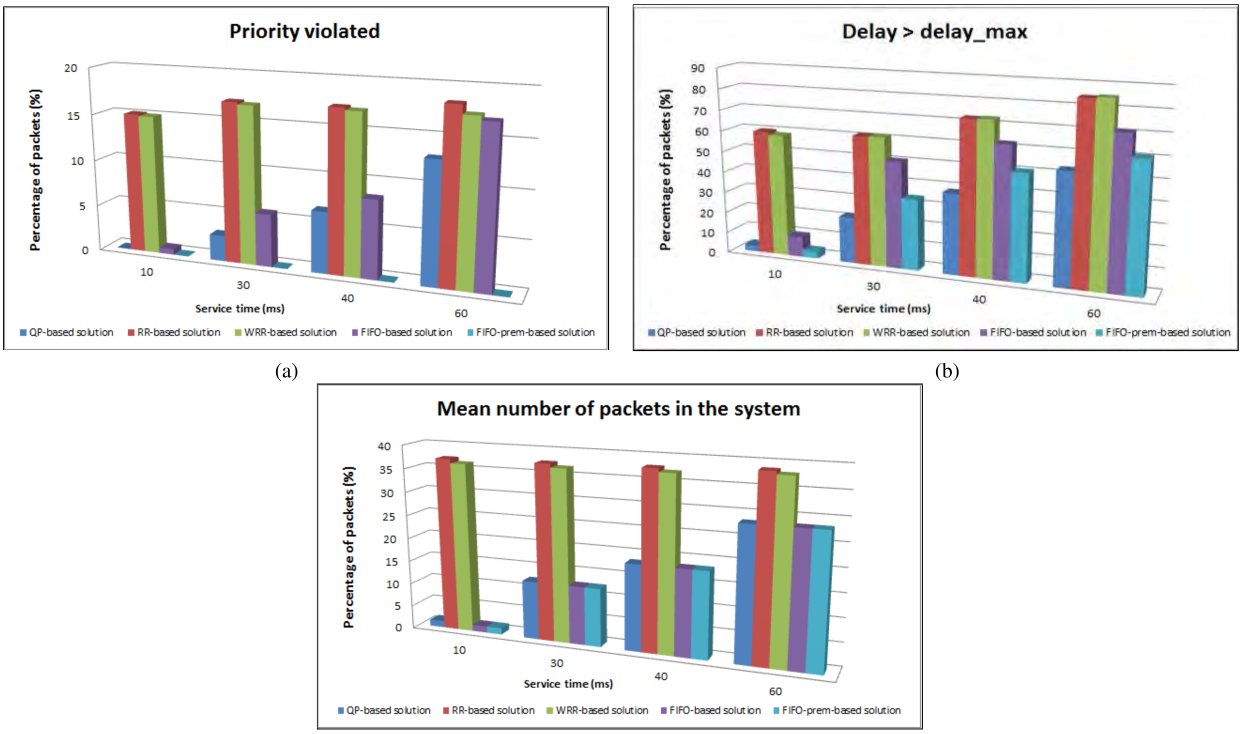


FIGURE 7. QP-SH performances compared to existing solutions (the service time is fixed to 30 ms).

zero packets for arrival rates more than 40 packets/ms. However, varying the arrival rate has no impact on the performance of RR and WRR based solutions since they mainly focus on providing a level of fairness between packets from different QoS levels.

In Fig. 7, we consider the first scenario where the service time is fixed and we compare the performance of our algorithm QP-SH and the existing RR, WRR, FIFO, and FIFO-prem based solutions. The comparison is made based on the percentage of packets that exceed their maximum deadline, the percentage of packets that do not respect their



(a) (b)

(c)

FIGURE 8. QP-SH performances for different values of service time compared to existing solutions. (a) Percentage of packets that exceed their maximum delays (%). (b) Percentage of packets that do not respect their priorities (%). (c) Mean number of packets in the system.

priorities, and the mean number of packets in the system for different values of arrival rates.

We note that the proposed QP-SH algorithm outperforms the existing solutions for the majority of criteria, with 15% higher for priority, 40% higher for the delay and 25% higher for the mean number of packets in the system. On the other hand, FIFO-prem based solution remains the optimal solution that guarantees priority criterion while increasing the delay since it is based only on priority. WRR and RR based solutions provide certain fairness between different QoS based packets while introducing the highest delay and the highest mean number of packets in the system. We also study the performance of the proposed QP-SH and the existing based solutions (Fig. 8) regarding the impact of varying the service time on a) priority violation, b) deadline violation, and c) mean the number of packets in the system. We note that when we increase the service time per packet, the performance of all solutions decreases and QP-SH maintains the lowest values except for FIFO-prem in priority criterion.

In Fig. 9, we compare the performance of our algorithm QP-SH in terms of the percentage of packets that exceed their maximum deadline, the percentage of packets that do not respect their priorities, and the mean number of packets in the system for different numbers of servers c and a fixed service time (30 ms). We note that our system works as well with many servers. Furthermore, the performance of the system increases when we increase the number of servers, with an improvement of 2% for priority, 10% for the delay and 6% for the mean number of packets in the system.

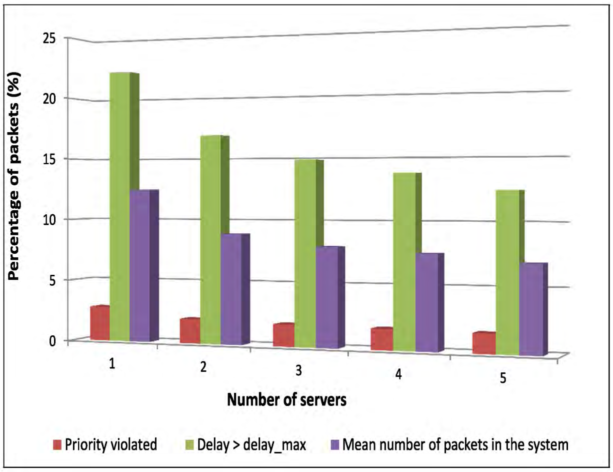


FIGURE 9. QP-SH performances for different numbers of servers (the service time is fixed to 30 ms).

VII. CONCLUSION

In this paper, we proposed a new dynamic queuing model for smart home network traffic generated by heterogeneous sources, to increase the number of packets that meet their deadline while preserving their degree of criticality. We tested our solution with 1000 network packets generated with different distributions. Then, we compared it to the existing based scheduling solutions for each criterion. Our experimental results demonstrated that the proposed algorithm outperforms the current solutions against almost all criteria.

REFERENCES

[1] D. Marikyan, S. Papagiannidis, and E. Alamanos, "A systematic review of the smart home literature: A user perspective," *Technol. Forecasting Social Change*, vol. 138, pp. 139–154, Jan. 2019.

[2] H. Pfeffer, "Dynamic changing tier service on test device," U.S. Patent 9 425 977, Aug. 23, 2016.

[3] (2018). *Internet Connection and Recommended Encoding Settings*. Accessed: Dec. 16, 2018 [Online]. Available: https://support.video.ibm.com/hc/en-us/articles/207852117-Internet-connection-and-recommended-encoding-settings

[4] W. Leister, P. A. Floor, Y. B. Woldegiorgis, I. Balasingham, and H. Abie, "Defining the asset lab," Norsk Regnesentral, Norway, Tech. Rep., 2012. [Online]. Available: http://publications.nr.no/1362402875/DART-16-2012.pdf

[5] J. Zhang, "Application-aware traffic prediction and user-aware quality-of-experience measurement in smart network," Ph.D. dissertation, Dept. Elect. Eng., Univ. Dayton, Dayton, OH, USA, 2018.

[6] P. Zhang, S. Wu, M. Wang, H. Yao, and Y. Liu, "Topology based reliable virtual network embedding from a QoE perspective," *China Commun.*, vol. 15, no. 10, pp. 38–50, Oct. 2018.

[7] M. Li, P. H. Tan, S. Sun, and Y. H. Chew, "QoE-aware scheduling for video streaming in 802.11n/ac-based high user density networks," in *Proc. IEEE 83rd Veh. Technol. Conf. (VTC Spring)*, May 2016, pp. 1–5.

[8] U. S. Premarathne, I. Khalil, and M. Atiquzzaman, "Reliable delay-sensitive spectrum handoff management for re-entrant secondary users," *Ad Hoc Netw.*, vol. 66, pp. 85–94, Nov. 2017.

[9] I. Benacer, F.-R. Boyer, and Y. Savaria, "HPQ: A high capacity hybrid priority queue architecture for high-speed network switches," in *Proc. 16th IEEE Int. New Circuits Syst. Conf. (NEWCAS)*, Jun. 2018, pp. 229–233.

[10] S. S. H. Shakir and A. Rajesh, "Performance analysis of two level calendar disc scheduling in LTE advanced system with carrier aggregation," *Wireless Pers. Commun.*, vol. 95, no. 3, pp. 2855–2871, Aug. 2017. doi: 10.1007/s11277-017-3967-z.

[11] A. Anand and G. de Veciana, "Measurement-based scheduler for multi-class QoE optimization in wireless networks," in *Proc. IEEE INFOCOM Conf. Comput. Commun.*, May 2017, pp. 1–9.

[12] T. Bakhshi and B. Ghita, "User-centric traffic optimization in residential software defined networks," in *Proc. IEEE 23rd Int. Conf. Telecommun. (ICT)*, May 2016, pp. 1–6.

[13] I. N. Bozkurt and T. Benson, "Contextual router: Advancing experience oriented networking to the home," in *Proc. ACM Symp. SDN Res.*, 2016, p. 15.

[14] H. Yang, X. Wang, C.-T. Nguyen, and S. Lu, "Optimizing user experience through implicit content-aware network service in the home environment," in *Proc. ACM Conf. Supporting Groupwork*, 2018, pp. 51–60.

[15] R. M. Abuteir, A. Fladenmuller, and O. Fourmaux, "An SDN approach to adaptive video streaming in wireless home networks," in *Proc. IEEE Int. Wireless Commun. Mobile Comput. Conf. (IWCMC)*, Sep. 2016, pp. 321–326.

[16] M. Zeng, S. Leng, Y. Zhang, and J. He, "QoE-aware power management in vehicle-to-grid networks: A matching-theoretic approach," *IEEE Trans. Smart Grid*, vol. 9, no. 4, pp. 2468–2477, Jul. 2018.

[17] M. M. Butt, E. A. Jorswieck, and A. Mohamed, "Energy and bursty packet loss tradeoff over fading channels: A system-level model," *IEEE Syst. J.*, vol. 12, no. 1, pp. 527–538, Mar. 2018.

[18] X. Zheng, Z. Cai, J. Li, and H. Gao, "A study on application-aware scheduling in wireless networks," *IEEE Trans. Mobile Comput.*, vol. 16, no. 7, pp. 1787–1801, Jul. 2017.

[19] S. Chaabnia and A. Meddeb, "Slicing aware QoS/QoE in software defined smart home network," in *Proc. NOMS IEEE/IFIP Netw. Oper. Manage. Symp.*, Apr. 2018, pp. 1–5.

[20] P.-C. Hsieh and I.-H. Hou, "Heavy-traffic analysis of QoE optimality for on-demand video streams over fading channels," *IEEE/ACM Trans. Netw.*, vol. 26, no. 4, pp. 1768–1781, Aug. 2018.

[21] S. Fan and H. Zhao, "Delay-based cross-layer QoS scheme for video streaming in wireless ad hoc networks," *China Commun.*, vol. 15, no. 9, pp. 215–234, Sep. 2018.

[22] N. Simon and R. Kavitha, "A study on smart home network architecture and technologies," *Int. J. Adv. Res. Comput. Sci.*, vol. 8, no. 3, pp. 74–78, 2017.

[23] M. Curado *et al.*, "Internet of Things," in *Cyber Resilience of Systems and Networks*. Cham, Switzerland: Springer, 2019, pp. 381–401.

[24] C. Gomez, S. Chessa, A. Fleury, G. Roussos, and D. Preuveneers, "Internet of Things for enabling smart environments: A technology-centric perspective," *J. Ambient Intell. Smart Environ.*, vol. 11, no. 1, pp. 23–43, 2019.

[25] P. Nowook, J.-H. Kim, H.-J. Park, and I.-T. Jeong, "Home network signal relay device in access network and home network signal relay method in access network using same," U.S. Patent 9 912 493, Mar. 6, 2018.

[26] M. Ben Attia, K.-K. Nguyen, and M. Cheriet, "QoS-aware software-defined routing in smart community network," *Comput. Netw.*, vol. 147, pp. 221–235, Dec. 2018.

[27] J. Xiao, S. Chen, and M. Sui, "The strategy of path determination and traffic scheduling in private campus networks based on SDN," *Peer-Peer Netw. Appl.*, vol. 12, no. 2, pp. 430–439, 2019.

[28] M. Haldas. (2018). *CCTV Camera Recording Video Frame Rate Comparison*. Accessed: Dec. 4, 2018. [Online]. Available: https://www.cctvcamerapros.com/CCTV-Video-Frame-Rate-Comparison-s/739.htm

[29] R. Li, Z. Liu, Y. Zhang, Y. Li, and Z. Fu, "Noise-level estimation based detection of motion-compensated frame interpolation in video sequences," *Multimedia Tools Appl.*, vol. 77, no. 1, pp. 663–688, 2018.

[30] S.-Y. Jeong *et al.*, "Method and apparatus for adaptive encoding and decoding based on image quality," U.S. Patent 15 287 345, Apr. 13, 2017.

[31] N. B. Shah, N. D. Shah, J. Bhatia, and H. Trivedi, "Profiling-based effective resource utilization in cloud environment using divide and conquer method," in *Information and Communication Technology for Competitive Strategies*. Singapore: Springer, 2019, pp. 495–508.

[32] R. Sharma, N. Kumar, N. B. Gowda, and T. Srinivas, "Packet scheduling scheme to guarantee QoS in Internet of Things," *Wireless Pers. Commun.*, vol. 100, no. 2, pp. 557–569, 2018.

MAROUA BEN ATTIA received the B.Eng. degree (Hons.) in telecommunications engineering from the National Institute of Applied Sciences and Technology, Tunisia, in 2013, and the M.Sc. degree (Hons.) in information technology engineering from the École de Technologie supérieure (ÉTS), Montreal, QC, Canada, in 2015, where she is currently pursuing the Ph.D. degree with the Computer Science and Engineering Department under the supervision of Dr. M. Cheriet. Since 2015, she has been with the Synchromedia Laboratory for Multimedia Communication in Telepresence, Montreal. Her research interests include routing, quality of service, and network management in the smart community networks. Her Ph.D. project has a real impact on the community of the Montreal region. Its main objective is to determine how the smart city, especially the smart home, can be mobilized as a vector of sustainable development. This involves establishing a robust, scalable, and reliable home network architecture that ensures a high quality of service and low resource consumption in the smart home. The outcome of this project will be used directly as a research prototype to develop a sustainable, smart city model for the Montreal innovation district on a larger scale. This first smart urban model is helping to advance the ICT sector in Canada and Quebec, in harmony with smart cities around the world.

KIM-KHOA NGUYEN received the Ph.D. degree in electrical and computer engineering from Concordia University. He served as the CTO of Inocybe Technologies, a leading company in software-defined networking solutions. He was an Architect of the Canarie's GreenStar Network and also involved in establishing CSA/IEEE standards for green ICT. He has led research and development in large-scale projects at Ericsson, Ciena, Telus, and InterDigital. He is currently an Associate Professor with the Department of Electrical Engineering, École de technologie supérieure, University of Quebec, Montreal, QC, Canada. He has authored 60 publications and holds several industrial patents. His research interests include network optimization, cloud computing, the IoT, big data, machine learning, smart city, high-speed networks, and green ICT. He was a recipient of the Microsoft Azure Global IoT Contest Award, in 2017, and the Ciena's Aspirational Prize, in 2018.

MOHAMED CHERIET (SM'95) received the M.Sc. and Ph.D. degrees in computer science from the University of Pierre et Marie Curie (Paris VI), Paris, France, in 1985 and 1988, respectively. Since 1992, he has been a Professor with the Automation Engineering Department, École de Technologie Supérieure (ETS), University of Quebec, Montreal, QC, Canada, where he was a Full Professor, in 1998. Since 1998, he has been the Founder and the Director of Synchromedia Laboratory, ETS, which targets multimedia communication in telepresence applications. He also co-founded the Laboratory for Imagery, Vision and Artificial Intelligence, ÉTS, where he was the Director, from 2000 to 2006. He is an expert in computational intelligence, pattern recognition, mathematical modeling for image processing, cognitive and machine learning approaches, and perception. His research has acquired extensive experience in cloud computing and network virtualization. He has authored or coauthored more than 350 technical papers in the field. He has coauthored a book *Character Recognition Systems: A Guide for Students and Practitioners* (Wiley, Spring 2007). He is a Fellow of the International Association for Pattern Recognition and the Canadian Academy of Engineering, the Founder and the former Chair of the IEEE Montreal Chapter of Computational Intelligent Systems, the Steering Committee Member of the IEEE Green ICT Initiative, and the Chair of the IEEE ICT Emissions Working Group. He was a recipient of the 2016 IEEE J. M. Ham Outstanding Engineering Educator Award and the 2012 Queen Elizabeth II Diamond Jubilee Medal. He serves on the editorial boards of several renowned journals and international conferences.

• • •