# Markov Chain based Packet Scheduling in Wireless Heterogeneous Networks 

Wahida Ali Mansouri ${ }^{1, \dagger}$, Salwa Hamda Othman ${ }^{1, \dagger}$ and Somia Asklany ${ }^{1, \dagger \dagger}$<br>${ }^{1}$ Faculty of sciences and arts, Turaif, Northern border University, Arar 91431, Kingdom of Saudi Arabia<br>${ }^{2}$ LETI laboratory, University of Sfax, Tunisia wahidamansouri1@yahoo.fr<br>${ }^{\dagger}$ Modern academy for science and technology, Maadi, Cairo, Egypt<br>Wahida.Smari@nbu.edu.sa


#### Abstract

Summary Supporting real-time flows with delay and throughput constraints is an important challenge for future wireless networks. In this paper, we develop an optimal scheduling scheme to optimally choose the packets to transmit. The optimal transmission strategy is based on an observable Markov decision process. The novelty of the work focuses on a priority-based probabilistic packet scheduling strategy for efficient packet transmission. This helps in providing guaranteed services to real time traffic in Heterogeneous Wireless Networks. The proposed scheduling mechanism is able to optimize the desired performance. The proposed scheduler improves the overall end-to-end delay, decreases the packet loss ratio, and reduces blocking probability even in the case of congested network.


## Key words:

optimization, Markov chain, scheduling, real time traffic, Heterogeneous Wireless Networks

## 1. Introduction

Heterogeneous Wireless Networks are basically composed of existing distinct Radio Access Technologies (RATs) like (WLAN, WiMAX, LTE-A, etc.). HWNs are useful in introducing effective management and scheduling of packets. The different RATs must coexist and interoperate together. In Heterogeneous Wireless Networks, different modules are required such as mobility management and session management, scheduling and admission control. Heterogeneous Wireless Networks consist of wireless transmission capable nodes which receives exogenous demand in form of packets. The nodes communicate these packets through a shared wireless medium. Hence their simultaneous transmission may contend with each other. The purpose of a scheduling algorithm is to give a priority to each packet.

Scheduling algorithms for Heterogeneous Wireless Networks need to be selected based on the type of traffic and their Quality of services ( QoS ) requirements. For real-time traffic such as voice, video and audio streaming, the most important QoS requirements are jitter, delay and loss rate. The QoS has been defined as a most important objective in Wireless Heterogeneous Networks. Therein, MAC layer
scheduling presents a challenging issue that seeks effective solutions to conform with the evolution of the different data traffic.
The most existing scheduling algorithms are based on one or two criteria or use classical scheduling. However, these algorithms are not dynamic and incur a large processing overhead and data transmission delay. In [1], an underlying information-theoretic principle is combined with a queuingtheoretic approach to achieve the guaranteed QoS. The authors propose a mathematical approach based on constrained Markov decision process to maximize the longterm average SP's revenue subject to long-term average queue length constraint.

In this paper, we develop an optimal scheduling scheme to optimally choose the packets to transmit. The optimal transmission strategy is based on an observable Markov decision process that provides guaranteed services to real time traffic.

The paper is organized as follows: in section I, the introduction is introduced. In section II, the related work of this research is summarized. Section III illustrated the proposed system followed by simulation results in section IV. Finally, the conclusion and the future scope of the presented scheduling algorithm are discussed in section V .

## 2. Literature review

In literatures, many researchers have proposed various packet scheduling schemes for providing better QoS support to the system based on the Markov chain. In research [2], Tian and al. address the problem of streaming video over wireless channels with error-prone feedback. They propose an optimal packet scheduling framework based on a partially observable Markov decision process. In work [3], a mathematical model based on Markov Chain is developed. Chowdhury and al. focus on the integration of call admission control and uplink packet scheduling mechanism to identify quantitative measurement of some QoS parameters. Research [4], presents a joint packet scheduling and dynamic bandwidth allocation scheme is proposed to
provide service differentiation and preferential treatment to delay sensitive traffic. The scheduler focuses on reducing the waiting time of high priority delay sensitive services in the queue and simultaneously keeping the waiting time of other services within tolerable limits. The work in [5] addresses the problem of multiuser scheduling in a cellular downlink system with partial channel information. The authors consider a downlink system with the channel of each user modeled by a two state Markov chain and demonstrate that ARQ feedback can be used to make informed multiuser scheduling decisions.

Several works have been done on opportunistic scheduling in a Markov modeled channel, e.g., [6] and [7]. It is understandable that the availability of the channel state information at the scheduler is crucial for the success of the opportunistic scheduling schemes. In work [8], Janowski and al. modeled and solved the system with two traffic classes and Earliest Deadline First scheduler with deadline values following the exponential distribution.
Recent research like [9] has proposed a new heterogeneous Internet of Thing (IoT) model based on the 3D Markov queuing model according to the user's priority. The proposed Markov chain calculates the length of the different queues of data based on the different priorities. [10] has proposed a novel fuzzy decision packet scheduling algorithm based on different QoS parameters. The authors introduced packet delay, channel quality, type of calls and of services as inputs to the system in orders decide the output Priority index. In [11], the authors consider a wireless broadcast network. They develop a structural scheduling algorithm and an index scheduling algorithm, leveraging Markov decision process (MDP) techniques. They applied the Whittle index for scheduling random arrivals.
[12] proposed a delay minimization scheduling problem with ergodic Markov channels in wireless networks. The proposed Markov decision process minimizes the expected transmission delay of each packet. [13] designed a packet scheduling method based on the Service Priority Dynamic Adjustment which dynamically regulates the power service priority on the node. Another work presented in [14] proposed a dynamic packet scheduling method. The authors estimate the packets sent on other paths based on two parameters which are the bandwidth and packet loss.

## 3. Proposed system

In this work, we consider two types of traffic streams: Real Time (RT) and Non Real Time (NRT). RT applications such as Voice over IP (VoIP) require a limited delay and cannot tolerate a delay greater than this limit. NRT applications are not time-demanding such as data traffic. To maintain a high level of QoS, RT applications need to be served by networks with minimal delay and packet loss while NRT applications need to be served at high speed.

We consider an area covered by different networks (RATs). Let $R$ be a set of RATs, $R=\{1,2, \ldots, i, \ldots N R\}$. We suppose that each $\operatorname{RAT}(i)(1 \leq i \leq N R)$ has a number of user of class $c$ called $N_{c}(i)$ where $c$ is a class of service in the RAT (i), $c=R T, N R T$ and a value of the requested SINR of class $c$ denoted $\operatorname{SINR}_{i, c, \text { req }}$. We define $\operatorname{SINR}_{i, c}(j)$ as the SINR of user $j$ in the class $c$ within $\operatorname{RAT}(i)$. The total number of users in the considered area is denoted $N$.
The comparison of the SINR of user $j$ in the different RATs in order to select a candidate RAT is not possible since the considered area is covered by a number $N R$ of heterogeneous wireless networks and each network transmits with different power levels and has a bandwidth, power consumption, signal reception strength and cost different from other networks. Remember that the mobile station is covered only if its SINR exceeds a certain threshold $\operatorname{SINR}_{i, c, \text { req }}$. We propose to compare the ratio $S_{i, c}(j)$ of the received SINR and the required $\operatorname{SINR} . S_{i, c}(j)$ is defined by the following equation :

$$
\begin{equation*}
S_{i, c}(j)=\frac{\operatorname{SINR}_{i, c}(j)}{\operatorname{SINR}_{i, c, r e q}} \tag{1}
\end{equation*}
$$

Here $\operatorname{SINR}_{i, c}(j)$ is the $\operatorname{SINR}$ of the user $j(1 \leq j \leq N)$ of the service class $c$ within the RAT $\mathrm{R}_{\mathrm{i}}(1 \leq i \leq N R)$ and the $S I N R_{i, c, \text { req }}$ denote the required SINR of class $c$ in the RAT ( $i$ ). If the ratio $S_{i, c}(j) \geq 1$, then the received SINR it satisfies the required SINR and therefore the user can have communication with the eNB or the access point of RAT $(i)$. If $S_{i, c}(j)<1$ then the user cannot communicate with this RAT.
To determine the capacity of heterogeneous wireless networks, we begin by calculating the number of users of each class of Service connected to each network (RAT). With the knowledge of the performance of the RAT (speed and coverage), the question which arises is to know which RAT should be chosen to serve the mobile station. We use the ratio $S_{i, c}(j)$ to determine the RAT of each user and then calculate the number of users.

### 2.1 Markov chain modeling

We model the system by a two-dimensional Markov chain $(X(t), Y(t))$ where $X(t)$ represents the number of RT calls and et $Y(t)$ represents the number of NRT calls at time $t$. The class RT calls (respectively NRT calls) arrive according to Poisson process with parameter $\lambda_{R T}$ (respectively $\lambda_{N R T}$ ) and exponentially distributed service time $\mu_{R T}$ (respectivement $\mu_{N R T}$ ) [15].
We define the state space $S$ containing all possible states in the system as follows:
$S=\left\{(n, m) \mid n \leq N_{R T}, m \leq N_{N R T}, n+m \leq N\right\}$
where $n$ (respectively $m$ ) represents the number of calls of type RT (respectively NRT) in the system at a given time. Let $q\left(n, m ; n^{\prime}, m^{\prime}\right)$ be the transition rate from the state $(n, m)$ the state $\left(n^{\prime}, m^{\prime}\right)$. Then, we consider the events describing the possible transitions as follows:

- $\quad q(n, m ; n+1, m)=\lambda_{R T}$ : call arrival or a handoff call arrival of RT class in RAT $(i)$.
- $q(n, m ; n-1, m)=n \mu_{R T} \quad$ : completion or departure of a call due to a failure of a RT vertical handover in $\operatorname{RAT}(i)$.
- $\quad q(n, m ; n, m+1)=\lambda_{N R T}$ : call arrival or a handoff call arrival of NRT class in RAT( $i$ ).
- $q(n, m ; n, m-1)=m \mu_{N R T}$ : completion or departure of a call due to a failure of a NRT vertical handover in RAT( $i$ ).


Fig. 1. State transitions Diagram
The Markov chain illustrated by the state transitions diagram in Fig. 1, presents a continuous-time Markov chain and a birth and death process. RT and NRT calls are independent. The stationary distributions can be written as follows:

$$
\begin{align*}
& \pi_{n}=\pi_{n}^{t}=\lim _{t \rightarrow+\infty} P[X(t)=n]  \tag{3}\\
&=\frac{\left(\lambda_{R T}\right)^{n}}{n!\left(\mu_{R T}\right)^{n}} \\
& \text { and } \\
& \pi_{m}=\pi_{m}^{t}=\lim _{t \rightarrow+\infty} P[Y(t)=m]  \tag{4}\\
&=\frac{\left(\lambda_{N R T}\right)^{m}}{m!\left(\mu_{N R T}\right)^{m}}
\end{align*}
$$

Let $P(n, m)$ be the stationary probability of having $n$ real time (RT) calls et $m$ non real time NTR calls in the $\operatorname{RAT}(i)$. This probability is equal to:

$$
\begin{align*}
P(n, m) & =\pi_{0} \cdot \pi_{n} \cdot \pi_{m}  \tag{5}\\
& =\frac{P(0,0)}{n!m!} \cdot\left(\frac{\lambda_{R T}}{\mu_{R T}}\right)^{n} \cdot\left(\frac{\lambda_{N R T}}{\mu_{N R T}}\right)^{m},
\end{align*}
$$

where $P(0,0)$ is the stationary probability of the system at the initial state. From the normalization equation $\left(\sum_{n, m} P(n, m)=1\right)$, we obtain:
$P(0,0)=\left[\sum_{n=0}^{N_{R T}} \frac{1}{n!}\left(\frac{\lambda_{R T}}{\mu_{R T}}\right)^{n} \cdot \sum_{m=0}^{N_{N R T}} \frac{1}{m!}\left(\frac{\lambda_{N R T}}{\mu_{N R T}}\right)^{m}\right]^{-1}$
We determine the blocking probability of new calls in the system (i.e. the number of calls in progress is greater than the total number of calls). This probability is given by $P^{B}$ as follows:

$$
\begin{equation*}
P^{B}=\frac{P(0,0)}{N_{R T}!N_{N R T}!} \cdot\left(\frac{\lambda_{R T}}{\mu_{R T}}\right)^{N_{R T}} \cdot\left(\frac{\lambda_{N R T}}{\mu_{N R T}}\right)^{N_{N R T}} \tag{7}
\end{equation*}
$$

For each user, we determine a set of RATs $r$ for which the ratio $S_{i, c}(j)$ is maximum using the following relation:

$$
\begin{equation*}
r=\arg \max _{i \in R}\left(S_{i, c}(j)\right), S_{i, c}(j) \geq 1 \tag{8}
\end{equation*}
$$

The set of RATs $r$ can be made up of one or more RATs. If $r=\left\{r_{1}\right\}$ then we increment the number users in RAT $r_{1}$. If the set $r$ is composed of more than one element $r=$ $\left\{r_{1}, r_{2}, \ldots, r_{K}\right\}$ then we propose another criterion to select the best RAT of the set $r$. In this case, the user selects a RAT as the one that is physically nearest to $i t$. We find out the nearest RAT (the distance between the user and the eNodeB or Access Point).
here $d_{j}^{i}$ is the distance between user $j$ and eNB/AP of $\operatorname{RAT}(i)$

$$
\begin{equation*}
r_{\text {candidate }}=\arg \min _{i \in r}\left(d_{j}^{i}\right) \tag{9}
\end{equation*}
$$

where $i \in r$. The maximum number of users $N_{c}$ in each RAT can be determined by the steps in the following Algorithm I.

Algorithm I. Determination of the total number of users in each RAT

Step 1: for $j=1$, determine the ratio $S_{i, c}(j)$ of the received and the required SINR
Step 2: determine the set $r$ of RATs where the ratio $S_{i, c}(j)$ is maximum Eq. 8
Step 3: if $r=\left\{r_{1}\right\}$ then $N_{c}\left(r_{1}\right)=N_{c}\left(r_{1}\right)+1$. If the set $r$ is composed of more than one element then find out the candidate RAT using Eq. 9 and calculate the number of calls as follow:

$$
N_{c}(\text { candidate_RAT })=N_{c}(\text { candidate_RAT })+1
$$

Step 4: for $j=2, \ldots, N_{u}$ determine all the $N_{c}(i)$ using the method in Step 1, Step 2 and Step 3.

### 2.2 Scheduling scheme

Packet scheduling refers to the decision process used by the scheduler to select which packets should be serviced in priority. The scheduling will be based on network characteristics like bandwidth, packet arrival rate, deadline of packet and channel quality.

The architecture of our proposed system is shown in the Fig. 2. We begin by determining the inputs of the scheduler and that affect the node priority index which are:

- Expiry time
- Channel quality (based on the SINR)
- Type of call (Handover, New call)
- Class of service (real time, Non real time)


Fig. 2. The architecture of the proposed system
Packets have end-to-end delay called a deadline and noted as $D_{\max , n}$ where $n$ the packet $n$. This constraint must be
guaranteed. The packet should not be delayed and must leaves the network before the expiration of the router $j$ offered delay $D_{i, j}$ of the $\operatorname{RAT}(i)$ must be met. The local deadline of packet $n$ in an intermediate router $k$ is determined as follows:

$$
\begin{equation*}
d_{i, k}^{n}=\frac{D_{i, j}}{n_{i}}+\Delta_{i, k} \tag{10}
\end{equation*}
$$

Where $\Delta_{i, k}$ is the amount of time left (residual delay) calculated as follows:

$$
=\left\{\begin{array}{c}
\Delta_{i, k} \\
\begin{array}{c}
\text { for } k=1 \\
\frac{D_{i, j}}{n_{i}}-\left(A_{i, k}^{n}-A_{i, k-1}^{n}\right) \text { for } k>1
\end{array} \tag{11}
\end{array}\right.
$$

Then, We calculate the most important parameter in this scheduler which is the earliest expiry time of a packet n :

$$
\begin{equation*}
\delta_{i, k}^{n}=D_{\max , n}-\left(A_{i, k}^{n}-t_{i, n}\right)-W_{i, k}^{n} \tag{12}
\end{equation*}
$$

here $A_{i, k}^{n}$ is the time when the $n^{\text {th }}$ packet arrives at the router $k$ in the $\operatorname{RAT}(i)$ and $W_{i, k}^{n}$ the waiting time of the $n^{\text {th }}$ packet in the router $k$ and $t_{i, n}$ is the time when the $n^{\text {th }}$ packet arrives at the first router of its path.

The priority of the packet is determined by the scheduling algorithm where the very low priority index indicates that the packet has a very high priority, it so it should be directly scheduled. The priority is calculated as follows:

$$
\begin{equation*}
P_{i, k}^{n}=\frac{1}{\delta_{i, k}^{n}} * \theta_{i}^{k}(j) \tag{13}
\end{equation*}
$$

Where $\theta_{i}^{k}(j)$ is a weight assigned to each quality level of the channel. This weight reflects the current state of the channel in the calculation of the priority index as follows:

$$
\begin{align*}
& \theta_{i}^{k}(j)  \tag{14}\\
& =\left\{\begin{array}{cc}
1 & \text { if } \operatorname{SIN} R_{i, c}(j) \geq \gamma \\
0.5 & \text { if } \gamma>\operatorname{SINR}_{i, c}(j) \geq \operatorname{SINR}_{i, c, \text { req }} \\
0 & \text { if } \\
\operatorname{SINR}_{i, c}(j)<\operatorname{SIN} R_{i, c, r e q}
\end{array}\right.
\end{align*}
$$

where $\gamma$ is the threshold value of the SINR.
The proposed algorithm is illustrated in Algorithm II.
algorithm II. Proposed Scheduling Algorithm

1. for all packets n do
2. Determine the local deadline (Eq.10)
3. Calculate the earliest expiry time before deadline expires
(Eq.12)
4. Calculate the priority index of packet (Eq.13)
5. Check the available bandwidth
6. if the channel quality is good or medium then schedule the packet with the highest priority else store the packet identifier in the queue B end
7. check the end-to-end deadline, such packets are the dropped
8. end

## 4. Performance evaluation

In this subsection, we evaluate our proposed scheduling algorithm based on two-dimensional Markov chain by simulation. This simulation models a network of randomly distributed mobile nodes within a $1000 \times 1000$ meters area. Data traffic is considered as Non-real-time applications (NRT) which is not exigent in term of delay. Voice over Internet Protocol (VoIP) is an example of RealTime applications (RT) which requires a limited delay and cannot tolerate a delay higher than this limit.

### 4.1 Simulation parameters

We assume that the mean arrival rate of new calls follows a Poisson process with parameter $\lambda_{\text {RT }}$ for the VoIP service and $\lambda_{\text {NRT }}$ for the data service. Interactive users follow the www model with an average of 5 pages per www session and 30s reading time between pages. The different types of data have the different requirements for QoS in the heterogeneous networks. We classify the different types of data into three queues. In each queue, packets are sorted based on the priority index. Packets with the highest priority are scheduled first. Table I presents the simulation parameters which are selected based on popularly deployed cellular networks (LTE), WLANs and WMNs.

Table I. Simulation Parameters

|  | LTE | WMN | WLAN |
| :---: | :---: | :---: | :---: |
| Bandwidth | 10 MHz | 10 MHz | 22 MHz |
| Path loss exponent $\alpha$ | 3 | 3 | 3 |
| Data rate | 100 Mbps | 300 Mbps | 11 Mbps |
| Transmit power | 46 dBm | 20 dBm | 20 dBm |
| Radius | 500 m | 50 m | 50 m |
| $\mathrm{SINR}_{\text {req }}(\mathrm{RT}$ ) | $-4 \mathrm{~dB}$ | 12 dB | 5.5 dB |
| $\mathrm{SINR}_{\text {req }}(\mathrm{NRT})$ | -4 dB | 9 dB | 5.5 dB |
| Packet size (RT) | 120 bytes |  |  |
| Packet size (NRT) | 1500 bytes |  |  |
| $\lambda_{R T}$ : Call rate (RT) | 10 calls/s |  |  |
| $\lambda_{\text {NRT: }}$ Call rate (NRT) | 5 calls/s |  |  |
| Average call length (RT) | 180s |  |  |
| Average call length (NRT) | 600s |  |  |
| Movement speed | $1 \mathrm{~m} / \mathrm{s}$ |  |  |
| $\begin{aligned} & \hline \mathrm{T}_{\mathrm{RT}}: \text { Transmission } \\ & \operatorname{rate}(\mathrm{RT}) \\ & \hline \end{aligned}$ | 12.2 kbps |  |  |
| $\mathrm{T}_{\mathrm{NRT}}:$ Transmission rate (NRT) | 64 kbps |  |  |
| Handoff deadline | 32 ms |  |  |
| Simulation time | 30 minutes |  |  |

### 4.2 Performance evaluation metrics and results

We evaluate our algorithm based on three metrics: call blocking rate, packet loss rate, and average packet waiting time. The block probability for handover calls and new calls for the real-time traffic (RT) is shown in Fig. 3. In this figure, we compare our algorithm called MCPS (Markov Chain based Packet Scheduling) with a reference algorithm proposed in [16] called NSA (New Scheduling Algorithm) and the EDF (Earliest Deadline First) policy proposed in the work of [17]. Note that the call blocking ratio for RT increases with the increase of the number of users in the network. The blocking probabilities values achieved by our algorithm are slightly lower than that of the NSA algorithm. The difference between the values reached is of the order of 0.016 for the MCPS and 0.017 for the second solution for 3000 users. This gain can be explained by the fact of introducing the quality of the channel in the calculation of the priority index. The blocking of calls caused by the quality of the channel is reduced. On the other hand, the blocking rate for the EDF algorithm reaches values greater than the MCPS and NSA since EDF does not consider the quality of the channel. Other remarks can be drawn for this last mechanism, it is that the curves of the blocking rates of the new calls and of the handover calls are confused this is explained by the fact that the EDF algorithm processes the calls without differentiation of the type of calls (handover call or new call).
The blocking probabilities of RT new calls reach greater values than handover calls. This is because our algorithm assigns more priority to RT handover calls. This ratio is around $2 \%$ observed for a number of users equal to 3000 .


Fig. 3. Real-time calls blocking probability vs. number of users
Fig. 4 illustrates the blocking rates of new and handover NRT calls against the number of users in the network. Notice that the curves start with negligible values. Then these values start to gradually increase as the number of mobile users increases. Indeed, for 3000 users, the blocking probability of NRT traffic does not exceed 3\% for the MCPS
and NSA algorithms. In our MCPS algorithm, NRT packets are stored in queue D. These packets will be served when the queue S of RT traffic is empty, and the channel state is estimated to be good for transmissions to minimize packet loss.


Fig.4. Non Real-time calls blocking probability vs. number of users
The portion of packets lost due to their timeout can be used to assess the performance of a scheduling mechanism. As much as this part is small, the scheduler is adequate to serve TR traffic. Fig. 5 and Fig. 6 illustrate the packet loss rate of our MCPS algorithm compared to the NSA and EDF algorithm when varying the number of users in the networks. In Fig.5, values reached by the NSA and EDF are higher than the values of our algorithm which justifies that the packets loss in these algorithms is due to the expiration of their delay and the bad quality of the channel.


Fig. 5. Packet loss ratio for RT traffic
Fig. 6 illustrates the comparison results of NRT packets loss of our MCPS algorithm with the NSA and EDF. We note that the EDF algorithm does not prioritize NRT traffic which causes queue congestion.


Fig. 6. Packet loss ratio for NRT traffic
A good scheduling algorithm should ensure packet delays before expiration. In Fig. 7 shows the average waiting time based on the number of users. We compare our algorithm with the NSA and EDF algorithms.
Packets from users with medium channel conditions experience higher delays compared to users with good channel qualities but remain lower than packet from users with bad channel qualities. Therefore, serving RT packets with high priority and good channel quality minimizes the number of packets lost due to timeout. This is explained by the effect of including the quality of the channel in the calculation of the priority index to minimize packet loss.

We compare the average delay of our algorithm with the NSA algorithm. We note that the NSA takes precedence over the delay for our MCPS algorithm. This justifies the main role of our algorithm, which is to serve packets before their deadline expires with good channel conditions.


Fig. 7. Average waiting time for real-time applications
The scheduling decision depends on the quality of the channel and packet delay information such as the waiting time in router queues. Indeed, the average waiting time of RT applications in scheduling mechanisms does not exceed their limit ( 150 ms ) when varying the number of mobile
users in networks. This delay varies from 12 ms when the number of users is 100 to 75 ms for 3000 users.

## 5. Conclusion

In this article, we have proposed a new dynamic algorithm based on an interworking architecture between different technologies like WLAN, WMN and LTE. The scheduling decision considers different metrics such as the class of service of each packet, the type of connection (handover or new call) and the quality of the channel. Our simulation results show the advantage provided by this approach in terms of differentiation between real-time and non-real-time traffic and between the type of calls (handover or new call). It also shows that the proposed algorithm can have a better performance when the size of the network increases (scalability). These results prove the effectiveness of our proposed method which considers the different types of traffic and guarantees the requested QoS .

## Acknowledgments

The authors gratefully acknowledge the approval and support of this research study by the grant no.7672-SAT-2018-3-9-F from the Deanship of Scientific Research at Northern Border University, Arar, K.S.A.

## References

[1] S. Niafar, X. Tan, Danny H.K. Tsang, "Optimal Downlink Scheduling for Heterogeneous Traffic Types in LTE-A Based on MDP and Chance-Constrained Approaches", Mobile Netw App, pp.1-12, 17 September 2015
[2] D.Tian, X. Li, G. Al-Regib, Y. Altunbasak and J. R. Jackson "Optimal Packet Scheduling for Wireless Video Streaming with Error-Prone Feedback" Wireless Communications and Networking Conference, pp.1287-1292, 2004
[3] P. Chowdhury, I. S. Misra and S. K Sanyal "An Integrated Call Admission Control and Uplink Packet Scheduling Mechanism for QoS Evaluation of IEEE 802.16 BWA Networks", Canadian Journal on Multimedia and Wireless Networks Vol. 1, No. 3, pp.14-30, April 2010
[4] R. Sharma, N. Kumar, N. B Gowda, T. Srinivas "Probabilistic Prediction based Scheduling for Delay Sensitive Traffic in Internet of Things ", The 6th International Conference on Ambient Systems, Networks and Technologies (2015)
[5] S. Murugesan, P. Schniter and N. B. Shroff "Multiuser Scheduling in a Markov-modeled Downlink Environment". Forty-Sixth Annual Allerton Conference Allerton House, UIUC, Illinois, USA September 23-26, 2008
[6] S. Lu, V. Bharghavan, and R. Srikant, "Fair scheduling in wireless packet networks",. IEEE/ACM Transactions on Networking, vol.7, no. 4, pp. 473-489, Aug. 1999.
[7] Y. Cao and V. Li, "Scheduling algorithms in broadband wireless networks",. Proc. IEEE, vol. 89, no. 1, pp. 76-87, Jan. 2001
[8] R. Janowski, A. Janowski "On EDF scheduler with the exponential deadlines", Zeszyty Naukowe WWSI, No 15, Vol. 10, pp. 7-17, 2016
[9] H. Wu, X. Wen, Z. Lu and Y. Nie, "3D Markov chain based multi-priority path selection in the heterogeneous Internet of Things", KSII Transactions on Internet And Information Systems Vol. 13, No. 11, pp. 5276-5298, November 2019
[10] W. Mansouri, S. Asklany and S. Othman, "Priority based Fuzzy Decision Multi-RAT Scheduling Algorithm in Heterogeneous Wireless Networks", IJCSNS International Journal of Computer Science and Network Security, VOL. 19 No.1, pp.1-10, January 2019
[11] Y. Hsu, E. Modiano and L. Duan, "Scheduling Algorithms for Minimizing Age of Information in Wireless Broadcast Networks with Random Arrivals", IEEE Transactions on Mobile Computing, Vol. 19, No. 12, December 2020
[12] W. Zhan, H. Tang, J. Wang, "Delay Optimal Cross-Layer Scheduling Over Markov Channels with Power Constraint", IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB), pp.1-6, September 2020
[13] B. Hu and al., "A Packet Scheduling Method Based on Dynamic Adjustment of Service Priority for Electric Power Wireless Communication Network", Wireless Communications and Mobile Computing Vol. 2020, 13 pages
[14] G. Xie, H. Chen, L. Xie and K. Wang, "A Dynamic Packet Scheduling Method for Multipath TCP in Heterogeneous Wireless Networks," 2018 IEEE 18th International Conference on Communication Technology (ICCT), 2018, pp. 678-682, doi: 10.1109/ICCT.2018.8600179.
[15] W. Mansouri and al. "Capacity Analysis in the Next Generation Wireless Networks Based on SINR Constraints". Journal of Networks, Vol 10, No 01 (2015), 29-38, Feb 2015, doi:10.4304/jnw.10.01.29-38.
[16] W. Mansouri and al."New Scheduling Algorithm for Wireless Mesh Networks", The 2nd International Conference on Multimedia Computing and Systems (ICMCS'11), Marocco, pp.1-6, 7-9 Avril 2011.
[17] S. Baruah, S. Funk et J. Goossens, "Robustness Results Concerning EDF Scheduling upon Uniform Multiprocessors", IEEE Transactions on computers, Vol. 52, No. 9 pp. 1185-1195, 2003


Wahida MANSOURI was born in Tunisia in 1983. She received her Ph. D degrees in Computer System Engineering, engineer's and master's degrees from the National Engineering School of Sfax (University of Sfax, Tunisia) in 2010 and 2017, respectively. Currently, she is member of the Electronics and Technologies of Information Lab of ENIS. Her research interests are communication networks, quality of services (QoS) and resource managements in next generation networks. Her current activities include scheduling and QoS provisioning for the integrated cellular networks and wireless local area networks (WLANs).

Salwa Othmen was born in Sfax in 1986. She received her Ph.D in Computer System Engineering from the National Engineering School of Sfax (ENIS) in January 2017. She is a member of Research Groups of SFAX Laboratory of Electronics and Information Technology (LETI) Laboratory since 2010. She obtained a Master Degree in Computer Science, Dedicated Computer System New Technologies (NTSID) from ENIS.. Her research activity is in the domain of optimizing and securing the routing protocols in the multi-hop wireless networks (Ad Hoc, mesh network and cellular networks). She is currently an assistant professor at North Border University Saudi Arabia, where she teach different courses related to computer science.


Somia ASKLANY: Born in Egypt 1969, Assistant professor, Computer science department, Northern Border University King Saudi Arabia. Bs.c in Computer science and mathematics Minia University 1992, MSc in numerical modelling2006 and Ph.D in Expert systems Cairo University 2011.

