Technical University of Sofia

Faculty of Telecommunications

Department of Communication Networks

Major: Telecommunications

Thesis

Topic: V2I traffic modeling

Student: Atanas Georgiev Vlahov

Faculty number: 111215122

Supervisors: Prof. PhD Vladimir Poulkov

Asst. Prof. Dipl.-Ing. Dr.techn. Stefan Schwarz

Sofia, 2019

Content.

2. Random traffic share and different bandwidth size.

LIST OF ACRONYMS AND ABBREVIATIONS:

3GPP	Third-Generation Partnership Project
5G	Fifth Generation of mobile communications
ACS	Automatic Stability Control
AEB	Automatic Emergency Breaking
BS	Base Station
BSM	Blind Spot Monitoring
CSMA	Carrier-Sense Multiple Access
D2D	Device-to-Device
DCC	Decentralized Congestion Control
DSRC	Dedicated short-range communication
ESP	Electronic Stability Program
FCW	Forward Collision Warning
HMI	Human-machine interface
IoT	Internet of Things
ITS	Intelligent Transport Systems
LDWS	Lane Departure Warning System
LISP	Locator/Identifier Separation Protocol
LTE	Long Term Evolution
MAC	Media Access Control
MP-TCP	Multipath Transmission Control Protocol
OBU	Onboard Unit
OEM	Original Equipment Manufacturers
P2P	Peer-to-Peer
RSU	Road Side Units
SCTP	Stream Control Transmission Protocol
SIP	Session Initiation Protocol
TCS	Traction Control System
TDD	Time Division Duplex

TDMA	Time-Division Multiple Access
UMTS	Universal Mobile Telecommunication System
USDOT	Department of Transportation of the United States of America
V2I	Vehicle-to-Infrastructure
V2R	Vehicle-to-Roadside
V2V	Vehicle-to-Vehicle
V2X	Vehicle-to-Everything
VPKI	Vehicle Public Key Infrastructure
WAVE	Wireless Access in Vehicular Networks
WPANs	Wireless Personal Area Networks

List of figures:

Table list:

I. Introduction

Thanks to the application of intelligent systems, in recent years the transport field has been developed particularly and Intelligent Transport Systems (ITSs) have been gradually replacing the traditional transport systems. Many of the main transportation problems (i.e traffic accidents and traffic congestions) can be solved with the assistance of new technologies. In order to do so, ITSs should be able to efficiently exchange data and communicate with the vehicles and the surrounding environment. The information exchange systems should use proper communication protocols, such as LTE-V2V and IEE 802.11p, which are designed to support vehicular communication systems. For the information exchange, ITSs should use proper communication protocols, like LTE-V2V and IEEE 802.11p, which are intended to support vehicular communication systems. WAVE (Wireless Access in Vehicular Environments) has been introduced in IEEE 802.11p. It is designated to ease V2I and V2V communications by improving the cooperation and coordination between infrastructures and vehicles. Also, it's able to provide 6- 27Mbps data transfer speed for short distances (around 300m).

For vehicular communications, cellular technologies can be used as an alternative to IEEE 802.11p. Third-Generation Partnership Project (3GPP) sets the fundaments of such usage. Thanks to the usage of a different radio interface along with some core network improvements, the capacity, and speed in UMTS (Universal Mobile Telecommunication System) is increased (over 100 simultaneous voice calls and data speeds up to 14.4 Mbps). With the introduction of LTE (Long-Term Evolution), data speeds are drastically improved (around 300Mbps for downlink and 75 Mbps for uplink). Also, there is an improvement in the transmission range (up to 100km) and transfer latency (around 5ms).
With the introduction of the fifth generation of mobile cellular networks, it is expected to have data transfer speeds up to 1Gbs and a latency of 1ms which will solve one of the main problems in vehicular communications (low latency requirements).

Other technologies that are possible to be used in vehicular communications are ZigBee/IEEE 802.15.4 and Bluetooth. The main goal of these technologies is to extend the transmission range and reduce the connection times, by allowing proper work in conditions of high vehicular density and mobility. Road safety and preventing dangerous situations are one of the main purposes of these technologies. That's why one of the fields of research in vehicular communications aims to provide efficient communication models that can be used in different scenarios.

The main intention of this paper is to provide an insight into vehicular communications, more specifically into Vehicle-to-Infrastructure (V2I) communications. Its main purpose is to propose a traffic model for V2I communications in 5G networks, that distinguishes and prioritizes the delay-sensitive traffic and schedule it first. The implementation is done using TU Wien's 5G System Level Simulator.

II. Literature review

1. Vehicular communications

1.1 Vehicle-To-Vehicle communications

Vehicle-To-Vehicle communications represent wireless data transmission between vehicles. It is used to prevent accidents and traffic congestions, by allowing vehicles to share data about their speed and position with the ad-hoc mesh network. The network uses a decentralized system to realize the connection, and the connection can be realized through a fully or partially connected mesh topology. (Figure 1)

In a fully connected mesh topology, nodes are connected directly to each other, while in the partially connected topology some nodes are attached only to those with which they exchange frequently the most of the information. In this network topology, the nodes can exchange data with the nodes to which they are directly connected (one hop when it's fully connected) or they can choose one of the available paths to reach the desired destination (multi-hop, when it is partially connected). This topology makes the network more stable. In case of a temporary malfunction or collapse of the node, routers are able to choose an alternative path to reach all destinations. In the past, these networks had been very expensive and difficult to realize, because mesh networks had been only wired. Nowadays these problems are overcome by using wireless communications and WPANs (Wireless Personal Area Networks).

In these networks, nodes form a graph of arbitrary size (in the case of partially connected mesh network) and instead of relying on a base station, nodes are forwarding packets to each other. These nodes are able to move and reorganize themselves randomly, though the network topology is inconstant (it varies rapidly). In addition, these networks can be connected to the internet in order to provide additional services or they can operate alone.

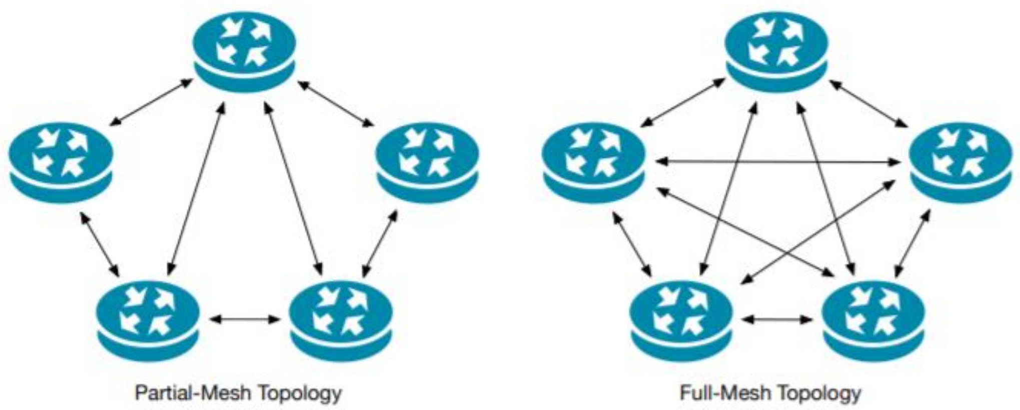


Figure 1. Mesh topologies [1]

Depending on the development of the system, it can warn the driver if there is a risk for an accident or the vehicle can interdependently take actions. Currently, there are embedded systems that are designed by the OEM (Original Equipment Manufacturers), but they are less effective than the V2V systems because OEM’s systems are depending entirely on the sensors and radars on board.

The V2V system actions are initiated according to specific parameters detected by these devices. Typically, the examined parameters are the distance from obstacles, travel speed and presence of other vehicles. Although the used technologies are very reliable, errors shouldn’t be underestimated. Contrariwise, V2V communications are used to improve security, by facilitating the interaction between vehicles. V2V protocols support the cars in dangerous situations in order to take action for solving the problem (i.e. driver’s fainting, system malfunctions, obstacles on the road and so on). The main task of each node of the mesh network is the data collection with which data it guarantees good security for the node and its neighbors (cooperative awareness). In order to accomplish an effective data access, a proper coding method should be selected. Some of the today widespread autonomous OEM (original equipment manufacturers) systems are:

- BSM (Blind Spot Monitoring): It monitors the blind spots of the vehicle. It uses radars integrated into the rear bumper to detect for approaching vehicles. It informs drivers for the presence of a vehicle in the blind spot by visual signal integrated into the appropriate door mirror.

- TCS (Traction Control System)/ ACS (Automatic Stability Control): Automatic Stability Control keeps the vehicle stable when there is not enough road grip and the stability and control of the vehicle is compromised. ACS automatically senses when there is a loss of traction and can be triggered automatically as soon as this is detected.

- ESP (Electronic Stability Program): This system acts when there are signs of a lateral heeling (i.e. due to oversteer or understeer movement). Its main purpose is to maintain the correct trajectory of the vehicle and to not allow the vehicle to leave the way.

- Forward Collision Warning (FCW): It is a radar sensor-based system that is monitoring the road. This system detects the distance between the vehicle and other objects on the road. If a risk of a collision is detected, the system alerts the driver.

- Automatic Emergency Breaking (AEB): It is a combination of advanced driver assistance and electronic stability control. It slows down the vehicle if a chance of collision is detected and reduce the severity of the impact when the collision is not avoidable. If the driver brakes inadequately or not at all, the pre-crash system acts in order to avoid or mitigate the impact.

- LDWS (Lane Departure Warning System): It is a device activated by a switch on the center console and warns the driver with an acoustic signal when passes the lane line without an apparent reason (i.e. without using the turn signal)

Thanks to the cooperation with these already existing OEMs V2V technologies can provide efficient management on the roadways. ITSs employ the data provided by these systems to improve the traffic management allowing the vehicles to communicate not only with other vehicles but also with the road infrastructures (i.e traffic lights or signs).

Before the full implementation of V2V communications becomes a fact, some obstacles have to be overcome: For example, common rules of security and operation should be accepted by all vehicle manufactures, confidentiality and privacy of the sent data should be guaranteed and it also should be decided whether public or private subjects should create and maintain the infrastructure. However in the past years many vehicle manufacturers have been focusing their research and development on V2V communications (i.e. General Motors, BMW, Audi, Daimler, and Volvo). As a result of this research, some experimental prototypes have been developed and USDOT (Department of Transportation of the United States of America) has documented the most advanced one. The documentation includes system requirements, algorithms, source codes and design documents of these prototypes. Some of the applications for connected vehicles can be seen in Table 1.

V2V Safety	Smart Roadside/Mobility	Agency Data/Environment
Emergency Electronic Brake Lights (EEBL)	Wireless Inspection	Probe-based Pavement Maintenance
Forward Collision Warning (FCW)	Smart Truck Parking	Probe-enabled Traffic Monitoring
Intersection Movement Assist (IMA)	Intelligent Traffic Signal System (I-SIG)	Vehicle Classification-based Traffic Studies
Left Turn Assist (LTA)	Signal Priority (transit, freight)	CV-enabled Turning Movement & Intersection Analysis
Blind Spot/Lane Change Warning	Cooperative Adaptive Cruise Control (CACC)	CV-enabled Origin-Destination Studies
Curve Speed Warning	Guidance for Emergency	Work Zone Traveler Information
Do Not Pass Warning (DNPW)	Emergency Communications and Evacuation (EVAC)	Dynamic Eco-Routing (light, vehicle, transit, freight)
Vehicle Turning	Connection Protection (T-CONNECT)	Low Emissions Zone Management
Bus Warning (transit)	Freight-Specific Dynamic Travel	Eco-ICM Decision Support System
Queue Warning (Q-WARN)	Emergency Vehicle Preemption (PREEMPT)	Eco-Smart Parking

Table 1. Applications for connected vehicles (arranged by USDOT)[4]

1.2 Vehicle to Infrastructure communication (V2I)

V2I is also mainly based on wireless technologies. In V2I interaction, vehicles can exchange data and connect to the Internet by fixed roadside infrastructure components (Road Side Units (RSU)). RSUs are base stations (BS) installed near the road (on the top of buildings, on traffic lights, bust stops, etc.). RSU is connected to the backbone IP network and communicates with the vehicle's OBU (Onboard Unit) in order to receive and transmit on-road information (i.e. accident warning, on-road traffic calculation) from and to nearby vehicles.

Due to its characteristics as a reliable way of communication, V2I communication is suitable for real-time applications. This technology is able to establish multi-hop communication routes in order to exchange data with other vehicles. It is used for on-road vehicular applications (i.e. safety and security, efficient usage of roads and crossroads, traffic congestion prevention and etc.). When referring to the communications between vehicles and ITS infrastructures, term Vehicle-to-Roadside (V2R) has also been used, but in this thesis, however, V2R and V2I are used interchangeably.

USDOT's Connected Vehicles Goals has summarized some of the main traffic safety goals for these systems, which you can see in Table 2. The main purpose of this communication is avoiding accidents, by exchanging safety and operational data between vehicles and road infrastructure. Alongside it can also provide a wide range of other safety, environmental and mobility benefits. V2I applies to a great variety of vehicles and road infrastructure and transforms them into "smart infrastructure".

ID	Category	USDOT CV Application
1	V2I/I2V Safety	Speed/Curve Speed Compliance
2		Oversize Vehicle Compliance
3		Speed Compliance/Work Zone
4		Red Light Violation Warning
5		Emergency Communications and Evacuation Information
6	V2I/I2V Pedestrian	Pedestrian in Signalized Crosswalk
7		Mobile Accessible Pedestrian Signal System (PED-SIG)
8	Mobility	Intelligent Traffic Signal System (I-SIGCVDATA)

Table 2. NYCDOT's Pilot Site Proposed CV Applications [2]

1.2.1 Architecture

Several V2I architectures are proposed, but all of them consist of the same key elements. According to the architecture of USDOT's, every V2I system should contain the following parts:

- Vehicle On-Board Unit or Equipment (OBU or OBE)
- Roadside Unit or Equipment (RSU or RSE)
- Safe Communication Channel

The OBUs are installed on the vehicle and they are the mobile part of the V2I system. Every OBU is divided into four parts (central control module, GPS module, human-machine interface (HMI) and wireless communication module). OBU accomplishes the communication between the vehicle and RSUs/ other vehicles. It frequently exchanges data with OBUs of other vehicles in order to support safety applications between the vehicles. Alongside it gathers the data needed to support public applications. It assembles the GPS and Vehicle data together into a series of snapshots and transmits it to the RSUs. Also it storages these snapshots of data and update/overwrite it after some period of time.

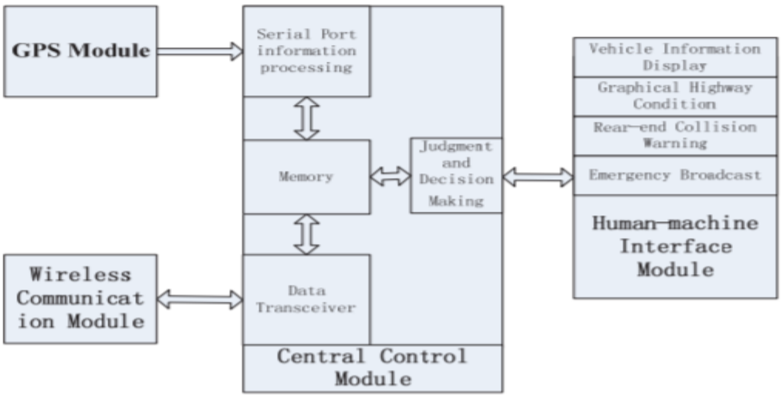


Figure 2. Global design diagram of the OBU [3]

RUS is typically mounted at intersections, interchanges and other locations (i.e. top of a building, petrol stations, road signs and etc.). RSU consists of a radio trans receiver (DRSC or WAVE), an interface to the V2I network, an application processor and an attached GPS unit. RSUs are connected to the V2I network and send private data to and from the OEMs through its interface to the V2I network. In order to ensure the bandwidth is not exceeded, RSU manages the prioritization of the messages send to and from the vehicles. For understandable reasons V2V safety applications have the highest priority, public and network applications have the lower priority and entertainment messages have the lowest priority.

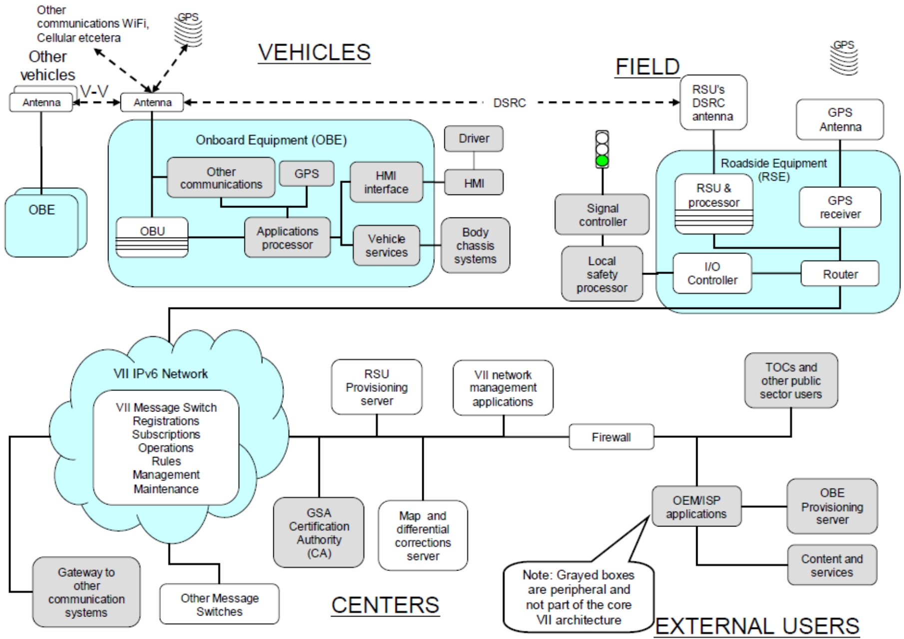


Figure 3. Architecture example of V2I systems. (Source: ITS Joint Program Office, USDOT)[5]

1.3 V2I communication standards.

Due to the standardization efforts in the field, we can see many standard protocols developed to support vehicular communication technologies. Dedicated short-range communication (DSRC) and Wireless Access in Vehicular Networks (WAVE) are typical examples of vehicular communication protocols. Another example of such protocol is LTE-V2X which was released at the end of 2016. The main purpose of these protocols is defining the communication architecture, frequency allocation, security algorithms and procedures, application management and etc.

1.3.1 **Dedicated Short Range Communication (DSRC)**

DSRC has been standardized and developed in the USA. It is based on IEEE 802.x and uses a 5.9 GHz licensed spectrum with seven channels with 10MHz bandwidth (from 5.850 to 5.925 GHz). The last 2 channels are reserved for special uses (service information and etc.), the one in the center is used as a control channel (CCH) and the rest are service channels (SCH). The control channel is used for safety applications, while the service channels are available for both safety and non-safety usage.

Some of the main functional attributes of DSRC are:

- Low latency: Delay during the opening and the closing connection is about 0.02s
- Limited interference: When it comes to radio interference, DSRC is very robust. Since it is a short range (~1000m), chances of interference from distant sources are very limited.
- Prioritization of safety application.
- Good performance during adverse weather conditions

1.3.2 Wireless Access in Vehicular Networks (WAVE)

In DSRC, the major modifications regarding the physical layer were made. The authentication processes were suppressed in order to improve the speed of network discovery and selection processes. However, there are multiple overheads inherited from the MAC layers, which makes it hard to provide fast data transfer, which is a must for vehicular networks. To solve this issue American Society for Testing and Materials, migrated the DSRC to IEEE802.11p (WAVE). WAVE integrates both the MAC and Physical layer and defines two types of devices: Onboard unit (OBU) and Roadside unit. OBU is installed in the vehicle and it’s used as a mobile device, while RSU is used as a stationary device. WAVE uses OFDM (Orthogonal Frequency Division Multiplexing) in order to divide the signal into narrowband channels. The physical layer and access media are based on IEEEE802.11a, while the channel planning and operational functions are based on IEEE 1609 standards. Resource management is defined by IEEE P1609.1, security is defined by IEEE P1609.2, network layer by IEEE P1609.3 and MAC is defined by IEEE P1609.4. Layer 2 (extension sub-layer) defines the interfaces between these applications and communication and stacks it with ASTM 2213-02 (IEEE 802.11a). Despite all of the work for improving WAVE functions, there are some challenges that need further research (i.e. latency and capacity requirements for safety applications, security and scalability and etc.)

1.3.3 **LTE-A for V2X**

Mobile networks are fast becoming the preferred technology for vehicular communication. LTE-D2D standard has laid the foundation of the LTE vehicular communications and 3GPP SA1 has started Rel-14 LTE-V2X at the beginning of 2015, with the support of a high number of companies. LTE provides superior network capacity, higher coverage, and greater mobility support. But on other hands in LTE latency is higher in comparison to 802.11p when the network load is bigger. Also, it requires the deployment of a small LTE base stations in order to ensure availability of real broadband for end users even in areas of high vehicle density. These small stations form microcells and they are deployed on the top of the existing LTE infrastructure. V2I communication is provided through LTE enabled OBUs or smartphones. An example of such OBUs is Qualcomm‘s Snapdragon X5 LTE modems for vehicles. One of the solutions to the mentioned above problems is the hybrid approach, which is combining LTE and 802.11p.

1.4 V2I applications

1.4.1. Safety-related Applications

Vehicular network applications include both safety- and nonsafety-related applications. Safety applications may include driver assistance and road hazards warning applications. The non-safety applications could be road traffic management, remote vehicle diagnostics, air pollution monitoring, and onboard comfort and entertainment. These applications are supported through V2V and/or V2I communications using various underlying wireless technologies. The V2I applications and related use case examples are discussed next.

These applications provide drivers of vehicles with information about different hazards and situations they generally cannot see. These applications include hard safety (time-critical) and soft safety (less time-critical) applications. While the former aims to avoid imminent hazards/crashes and minimize damages when crashes become unavoidable (mainly supported by V2V Communications), the later primarily enhances driver safety awareness without requiring any instantaneous reactions. In V2I safety-related applications, achieving the required low latency which generally varies from under 100ms, is the typical challenge. Thus, V2I communications are expected to support soft safety applications which include intelligent traffic signs, weather conditions, construction zones, and traffic congestion. The study has explored various cases of V2I road safety applications, which are summarized in the table below.

Type of Application	Brief description	Potential benefits
SWIW/OVW [22]	This application warns drivers about severe weather conditions that may impact travel conditions	The driver takes extra care to avoid any weather-related accidents
RSZW [22]	RSZW increases the vehicle's driver awareness about the posted speed limit in reduced speed zones and changed roadway configurations.	Avoids any possible accident for pedestrian or infrastructure destruction at work zones
RCVW [22]	RCVW is designed to warn drivers about railroad crossing locations	Avoids any possible crash between a vehicle and train
SVW [22]	SVW increases vehicle drivers' awareness about stop-sign locations	No Stop-sign violation

Table 3. Road safety applications [2]

1.4.2. Traffic Efficiency and Management Applications:

These important applications aim to improve the management of traffic on the roads by providing users with traffic assistance and by updating local traffic information. A vehicle or RSU collects information about traffic conditions, passes this information to other vehicles either directly or through a remote server, enabling them to choose an alternate route that optimizes the travel time. Speed management (e.g., Regulatory contextual speed limit notification and Green light optimal speed advisory) and co-operative navigation are the primary examples of traffic management applications. Typical actions in response to traffic management messages would be to proceed with caution or to take an alternative route to avoid the dangerous conditions ahead.

1.4.3. Entertainment and Personalized Applications:

These applications may help to extend internet access to the moving vehicle so that passengers may continue with their office or homework while on the move. Passengers are not only able to access different Web applications, but also applications relating to VoIP, video, multimedia streaming, navigation and localization services. Examples of such applications include searching for the nearest gas station or looking for the nearest McDonald‘s restaurant. These applications do not usually suffer from stringent communication constraints, like restricted latency or packet loss, although high data throughput may be occasionally required. The effectiveness of these applications depends on the capability of V2I communications‘ underlying network technologies. While the WAVE enabled RSUs to provide reduced latency for V2I communications, the development of cellular technologies has extended their capabilities to provide reduced latency for safety applications. The study has shown that the recent version of LTE provides low transmission latency, a very high data rate (> 100Mb/s) and it can tolerate high mobility. Both of these features are imperative for road safety applications.

2. 5G for Vehicular Networks

2.1. Vehicular Networking in 5G

Current vehicular communication architectures have some difficulties in fully meeting the application’s requirements, due to network congestions, mobility or other bottlenecks. To overcome this, some new researches in the field of wireless transmission, network architecture and protocols are required. These new researches are focused on reducing the serialization, queuing and network processing latencies. 5G technologies are the main candidates for improving vehicular communications. According to some of the main vehicle manufacturers (i.e BMW, Volvo, General Motors), capabilities that come with the 5G mobile networks are the solution for the full deployment of self-driving cars on streets. In this section, we are going to review and analyze some of the fundamental aspects of vehicular communications through 5G.

2.1.1. Network Congestions

Both of DSRC and LTE systems have limitations due to network congestions. Due to the limited available resources, entities are competing to access the shared radio resources which in turn leads to increased latency. In order to avoid network congestions resulted by ineffective network size or high traffic load, several solutions have been proposed:

- Congestions can be avoided by grouping a large number of users into clusters. By using clustering-based algorithms we seek for reduction of consumption of radio resources via integrating information, reusing radio resources in other clusters and working as one station.

-Contention level can also be reduced by reusing radio resources via the smart allocation of multiple non-overlapping channels, like multichannel MAC design.

- Having a technology that works even when the network operator is not there (i.e out of range) is critical for this type of communications. Device-to-Device (D2D) communications are the most likely solution for such situations. By using p2p communications D2D technologies reduce contention level in the system and at the same time it avoids redundancy in the signals for controlling and maintenance.

For DSRC systems, several DCC (Decentralized Congestion Control) mechanisms are proposed. Some of them work on the same principle as D2D, clustering and multichannel MAC design (i.e the transmit power control or the DCC sensitivity control mechanisms), while others look at congestions from the perspective of offered load (i.e. transmit data rate control, transmit access control mechanisms and the proposed TCP-like congestion control). It is less likely a single radio resource management scheme to be used in the future, but more like a combination of schemes. A novel design of the integration of the mentioned above congestion solutions will help to reduce the latency caused by congestions.

2.1.2. Mobility management

Contrary to the stationary IoT networks, nodes in vehicular networks are constantly moving on semi predictive but predefined trajectories over an underlying road topology. This mobility results in the need for network mobility because vehicles travel under the coverage of different RSUs. Mobility and sparse RSU deployment result in intermittent connections for DSRC networks, which in turn increases delays due to hand-off procedures.

In 5G this problem is envisioned too. Due to the densification of the network by using small cells, the probability of hand-off will be increased. Thus, many types of research are focused on optimizing hand-off procedures in order to reduce the delay and providing more reliable connections. Some solutions based on Mobile IP (MIPv6) and Network Mobility (NEMO) are aiming to overcome this challenge by utilizing predictive techniques to perform registration in advance. However, MIPv6's architecture is not efficient for highly dynamic networks and a distributed mobility management approach would be more appropriate to be used in vehicular networks.

Other solutions include LISP, SCTP, and Media Independent Handover. Location Identity Separation Protocol (LISP) supports mobility with route optimization, multi-homing dual-stack, and network mobility and it is compatible with ICN principles of locator/identifier separation. Stream Control Transmission Protocol's (SCTP) inherent multi-homing support and Media Independent Handover (IEEE 802.21 standard) have also been investigated for efficient handover.

In conclusion, a combination of minimal signaling and predictive techniques (know as "make before break") and efficient use of all available access networks would reduce the delay resulted from the user's mobility.

2.1.3. Backhaul network

In 5G mane, services have been moved to a cloud base architecture but in ITS, the location of the servers would impact the performance of the service, because control techniques can tolerate very low latency and content has to be available seamlessly. Currently, cloud-based services are centralized in data centers and the connection of the network provider with that data center impacts the delay. This is a major source of the latency for applications that require such connections. Also, the latency depends on the type of the links, the distance between the network gateway and the remote host and the routing rules as well. In order to reduce the backhaul delay, the remote host is brought closer to the end-user, this technique is known as fog computing. In vehicular networks, fog computing entitles spatially distributing service components and caching content within a flat network provider core and the vehicles.

2.1.4 Air Interface

As mentioned before, 1ms end-to-end latency is possible in congestion-free networks. A proposed division of the delay in different components is shown in Fig.4 Such architecture requires 0.2ms for the air interface, which accounts for 15 timeslots (according to IEEE 802.11p standard). Current access categories (ACs) with low priorities in the IEEE 802.11p cannot satisfy the delay limits. In theory, only AC0 and AC1 can fulfill this requirement and only in cases of very low contention level and no collisions. In current LTE systems, the duration of a single frame is 10 ms, consisted of 10 sub-frames with two slots per sub-frame. Large packets require segmentation to over multiple sub-frames, which results in an increase of the transmission delay for those packets. Also from these 10 sub-frames, a maximum of 6 can be used for MBMs, while the rest are reserved for unicast traffic. This reduces the system performance for broadcast packets, which is the majority of the traffic in ITS applications.

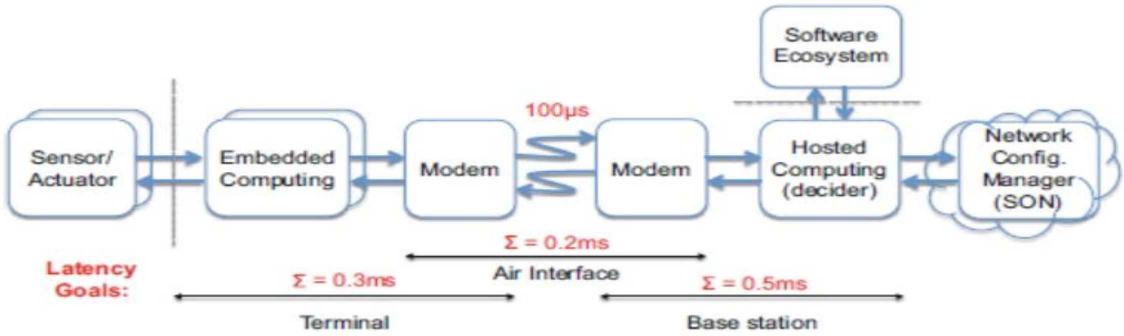


Figure 4. Proposed end-to-end delay budget in 5G [6]

In order to achieve a 0.2ms air-interface delay, relevant mechanisms should be redesigned. A mixture of TDMA and CSMA access in the super-frame cycle is one of the potential solutions for achieving such delay (i.e FlexRay). Ultra-high AC would have deterministic delay through the TDMA part of the frame cycle and lower AC's would content on the CSMA cycle. In METIS 2020 project several air-interface solutions for vehicular communications have been analyzed and proposed. A most notable one is the coded slotted Aloha (CSA)MAC technique. It enables a reliable ad-hoc communication that can operate in almost double the network sizes of IEEE 802.11p standard networks. The current LTE frame structure is not able to fulfill the delay requirements, but a flexible TDD frame structure is proposed as a solution. Its target is to provide an overall latency of 1ms in order to support future ITS cases (i.le autonomous driving, real-time control, and the tactile Internet.

2.1.5. Security

In vehicular communications, providing high security is really important. However, providing a secure network always results in higher processing and communication latency. In networks with higher density processing time drastically increases, and high beacon rate (e.g., 10Hz for CAMs) leads to significant computational overhead in terms of authentication. Around 43% of the delay due to security is spent on the processing/checking of the legitimation of the messages. Lightweight filter security schemes are the potential solution for reducing the overall processing time. In such techniques, the amount of the checked each time information is reduced and allows fewer checks according to the "credit" for each user.

Unfortunately, these techniques are not fast enough to provide a 1ms delay budget for the most demanding applications. In order to perform faster computation, physical layer security is consisted of SoC (system of the chip) and is exploiting propagation randomness to establish secret keys. Also, current VPKI which manages security keys and pseudonyms has to be redesigned since most of the V2V applications use broadcast messages. However in the 5G architecture, DSRC and cellular technologies are expected to be tightly coupled, so security protocols will be shared or at least compatible, which in turn will allow vehicles to communicate through any technologies.

2.2 Conceptual Vehicular Architecture in 5G

Due to the requirements for ultralow latency, high-reliability, and security, in 5G a revolutionary new vehicular communications architecture is proposed. It follows a revolutionary new approach, rather than simply upgrading the current 4G architecture.

The proposed architecture design is divided into four layers:

a) Cloud/internet service layer (known as Generic Cloud)
b) Core network cloud layer
c) Radio access network layer
d) Vehicle and RSU space

In the past few years cloud computing technologies have been drastically improved, so each of these layers is represented as a cloud. Also, the proposed ecosystem is designed to support current and future ITS applications in the most demanding scenarios. Each layer is described in more detail in the next sections.

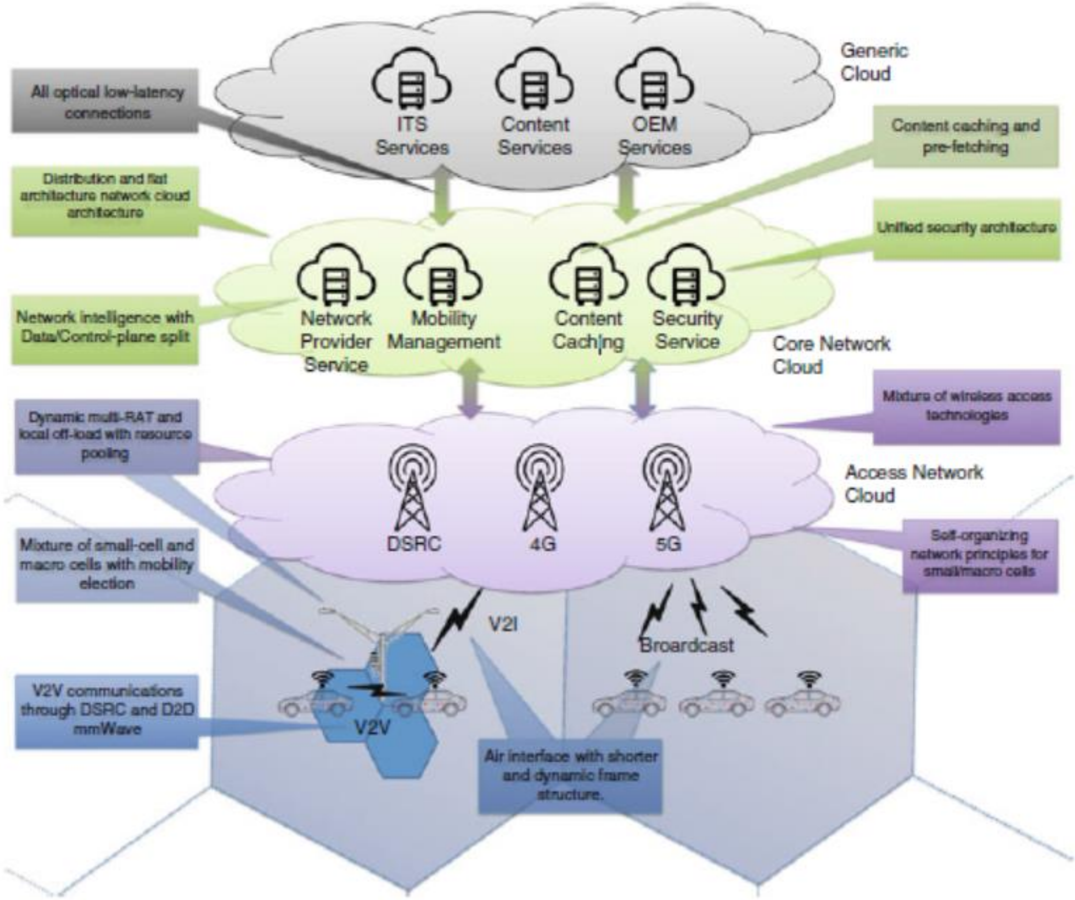


Figure 5. Conceptual 5G architecture [6]

2.2.1 Generic Cloud

The generic cloud consists of the current cloud infrastructure where ITS-specific services are provided. Generic content providers are included too. Also even in the generic cloud, services have to be spatially distributed and the connections between data centers and the core network have to be all-optical so it can provide low-latency and high network reliability.

2.2.2 Core Network Cloud

Current LTE core network architecture should be completely redesigned. Due to the reduction of network entities, supporting an all- IP network and smart off-loading functions, it is simpler than the previous 3G core architecture, but the delay is still significant. In the proposed 5G architecture most of the current EPC functionalities are represented by the network provider services (NPS). The implementation of those NPS is based on the principles of network function virtualization with distributed and flat architecture, which is forming clusters of services. Studies show that this will reduce the signaling traffic by 70% which will result in the reduction of end-to-end delays. Also, cluster-based architecture in the core network reflects on the access network cloud.

Since mobility is the main characteristic of the vehicular networks and it is one of the most important challenges for 5G networks, it has been on focus for many research activities.

Some of these researches propose to push the intelligence in the mobile terminal. This improves efficiency and scalability because the terminal can easier identify its own flow than the core network. A software-defined network approach that provides connectivity management as a service (CMaaS) is proposed too. It consists of high-level intelligence (4+ layers), which is maintaining network connectivity by using protocols such as MP-TCP, SCTP, and SIP. Also in CMaaS, in order to provide service differentiation at connectivity level, a hierarchical network control with different levels of complexity is proposed.

One of the main solutions for providing low latency is the ability to cache content within the network. 2/3 of the global mobile traffic consists of multimedia accounts. Pre-fetching and caching popular content in the intermediate nodes helps to provide this content to the user faster. Such an approach of caching can be implemented within the EPC (Evolved Packet Core) as well as the RAN (Radio Access Network). Depending on the amount of cached content, these techniques are able to reduce the delay by up to 80%. Problems related to congestion and latency can be solved also by functionality distribution throughout the core and access network. This is known as fog computing and it provides network services, computing, and storage at the network edge. It also adds intelligence in the network platform itself in order to organize the underlying resources in real-time.

Last component of the core network cloud is the security management. A sophisticated distributed VPKI (Vehicle Public Key Infrastructure) should be implemented in order to meet the delay and at the same time provide the needed security. With the increasing of the density of the radio access network and the usage of multiple radio access technologies, the probability of re-authentication and handover of the vehicle increases too. Again an SDN-based approach for authentication can be proposed as a solution. In this technique user-related security information (i.e physical layer characteristics, location and identity) is shared by the SDN controller to the predicted next cell access point.

2.2.3 Access Network Cloud

In the future 5G networks, the wireless access network will be also significantly redesigned. However, technologies such as 4G and DSRC will be used too. In the 5G access network cloud, several technologies will be able to work in synergy in a multitier architecture. Through mm-waves and DSRC, macrocells will be providing service coverage to D2D communications for vehicular communications. By implementing software-defined radio (SDR), hardware and software will be decoupled, which will allow network intelligence to control the available resources with higher flexibility from both the network and user perspectives. Also, operators can faster configure the network capacity, which facilitates network expansion and bottleneck elimination. But this results in an increase in operational complexity, so it needs an advanced self-organizing technique for the underlying networks. By using big data analytics of control, signaling and contextual data in the core network, an unprecedented amount of system-level intelligence is provided. This intelligence can be integrated into proactive self-organized network (SON) engines in order to provide better resource allocation. The dynamic multi-RAT access will also increase the performance and reduce the cost of communications by stimulating further off-load from macro to small cells or DSRC. When the user should switch from one technology to another is on focus for many research works. At the moment, measurements based on SNR are the preferred method. However, there are researches that have proposed the load aware user-centric schemes, which combine the SNR measurements and information about network loading in order to improve performance.

2.2.4. Vehicle Network Cloud

The last layer of this architecture consists of RSUs and vehicles. Fog computing can be implemented within this layer too. RSUs form a cloud employing SDN in order to dynamically initiate, migrate and/or replicate services. By using this technique, the reconfiguration costs and infrastructure delay are improved.

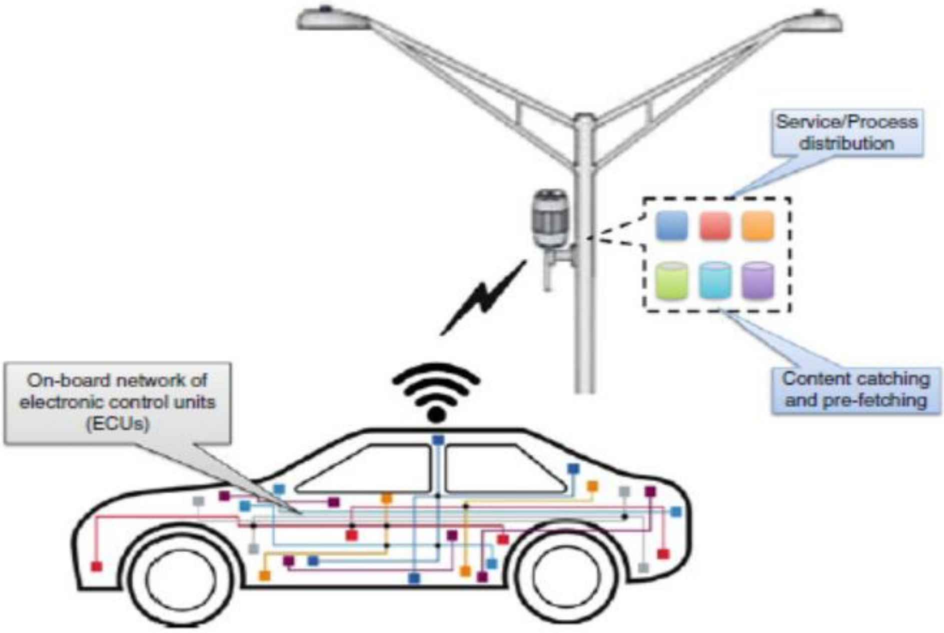


Figure 6. Vehicle and RSU cyber-physical ecosystem [6]

The number of electronic control units (ECUs) in the vehicle is increased. These ECUs along with the sensors and actuators are called Cyber-Physical System (CPS). Due to CPSs, the vehicles can be considered as a resource for sensing, computing, data storage, data relaying and a means for locating other objects. However, there is a trade-off between the QoS communications capabilities and sophisticated sensors and computational on-board. Less sophisticated onboard systems require more cloud-based assistance and higher QoS communications.

Also, current researches are focused on bringing "the vehicle on the internet", by converting the vehicles from a content consumer to a generator. This increases the need for high-speed links in uplink transmissions for both V2V and V2I communications. Also in the early stages of infrastructure deployment, vehicles could form a vehicular cloud (i.e parked vehicles can be used as a temporary network and storage infrastructure which can increase the resilience and reliability of the vehicular network).

3. TU- Wien 5G system level simulator

In the cellular communications, simulators are really important tools for a comprehensive study of the network and all of the players in it. For gaining insight of the performance of large-scale scenarios, a real world approach is too expensive and laborious, that’s why system level simulators are developed alongside with the standardization processes.

3.1. Simulation methodology

TU-Wien 5G SL Simulator allows the investigate of the performance for large-scale wireless networks. The implementation is done on Matlab by using OOP model. Performance evaluation is based on Monte-Carlo simulations with a large number of scenarios based and generated according to parameters specified before the simulation. A detailed description of a typical simulation will be discussed in next sections.

Individual link quality is defined by the geometry and the position of the receivers and transmitters, as well as several propagation effects. After that, the received power of all links is combined in a SINR value, which is used later in the transmission function.

Network elements such as BSs and Users can be defined and placed according to predefined placement functions (such as classical hex-grid or Poisson Point Process (PPP)).

The propagation effects which influence the quality of the transmission are split into different functions. Each option and model can be chosen independently per BS, user type and scenario. The simulator provides the following options:

- Individual transmission power for different tiers
- Antenna patterns
- Large scale path loss
- Small scale fading in term of channel models

The MAC layer is represented by the scheduling function and by applying Adaptive Modulation and Coding (AMC). By utilizing the calculated SINR of the active links it defines the resource allocation and the appropriate Modulation and Coding Scheme (MCS) needed for the transmission.

Transmission is abstracted, and several steps in the receiver and transmitter chains are combined in two functions, called Link Quality Model (LQM) and Link Performance Model (LPM). This abstraction allows large-scale simulations with thousands of different nodes in a single simulation.

After the simulation, the acquired results are stored and later can be used for plotting or processing.

2.3.Internetwroking with 5G systems

Next generation of cellular network systems are expected not only to handle the massive growth of the mobile data demand, but also to support a wide range of different wireless application. According to standards, 5G will lead to higher data rate and capacity, lower latency, higher reliability and better energy efficiency. In order to satisfy all of the requirements, due to novel heterogeneous network solutions vehicles will be able not only to consume/generate data, but to help the network in delivering the traffic.

2.3.1. Heterogeneous (HetNet) Vehicular Networks

Enabling higher user data rates and increasing the network capacity by using HetNet is a popular solution in cellular systems. In HetNet various cells with different coverage (macro, small and femto-cells) are coexisting. In the context of vehicular communications, Vehicular HetNets (Het-VNETs) represents a complex systems in which are integrated various radio technologies from DSRC (IEEE 802.11p) to cellular networks.For example, due to the reduced latency DSRC can be used for delivering data that is generated from safety and early warning applications, while LTE networks can help with sending multimedia traffic that demands high bandwidth. However there are some challenges for adoption of multi-radio technologies. Before the full implementation of Het-VNETs, following challenges should be overcame:

-Sophisticated interference cancellation techniques should be used in order to reduce the interference originated from nearby cells.

- Need of a dynamic support of multiple network configurations

- Need of a flexible and effective radio resource management strategy

- Need of Efficient mechanism to satisfy the QoS requirements, which are different for each type of vehicular application.

In Fig.7 is proposed an architecture for a 5G vehicular communication systems, which is able to perform the joint management of multiple radio technologies and the resource allocation according to the requirements for the specific vehicular application.

In this solution, the core network handles key functionalities by being directly connected to remote servers which are offering diverse services for the users. Also the RAN of this HetVNET is consisted of both eNBs and RSUs, and both of them are used for the V2I communication.

Also there are 2 possible communication modes:

1. V2I: It is performed through the cellular link through a BS or the DSRC interface by using the RSUs deployed near the road.

2.V2V: Direct communication between vehicles via LTE D2D connections or DSRC links.

Radio access technology adaptation is fully accomplished by integrating a VLL (Virtual Link Layer) within the core network protocol stack. With VLL the network can handle the different radio interface and the load balancing among the different systems. And it guarantees the QoS for the safety and entertainment applications.

The adaptation of radio access technology is fully accomplished by integrating a Virtual Link Layer (VLL) into the core network protocol stack. VLL includes all the functionalities required to handle different radio interfaces. It also handle the load balancing among the different systems. Also it helps to guarantee the QoS for safety and entertainment applications.

Also network virtualization methods are applied, in order to obtain virtual resources, which abstracts the radio resources and the physical layers of each wireless system used in the Het-VNET.

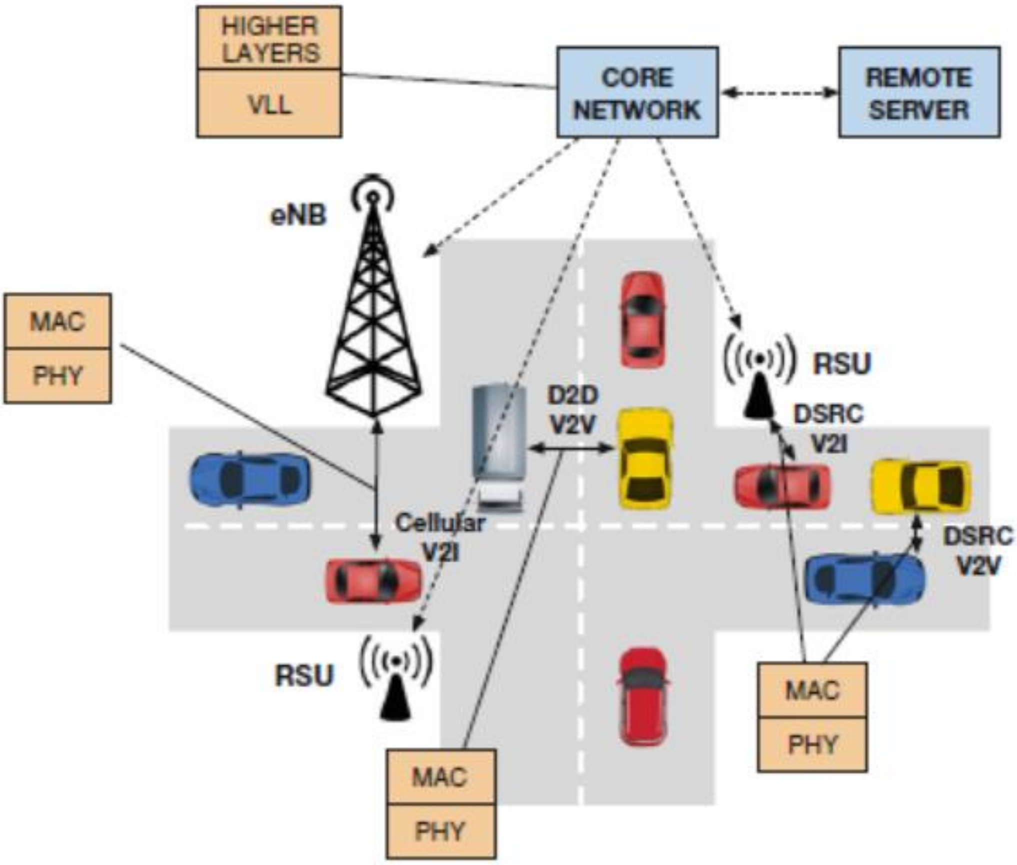


Figure 7. Proposed 5G architecture for the integration DSRC-LTE in vehicular applications [6]

2.3.3. Multihop D2D Paradigm

The first step toward a fully integrated networking envisaged by 5G systems is choosing an integrated framework which manages the network and services discovery, users scheduling and resource allocation, depending on the network configuration. The most popular proposed integrated design is consists of 3 phases, which are visualized on Fig. 8:

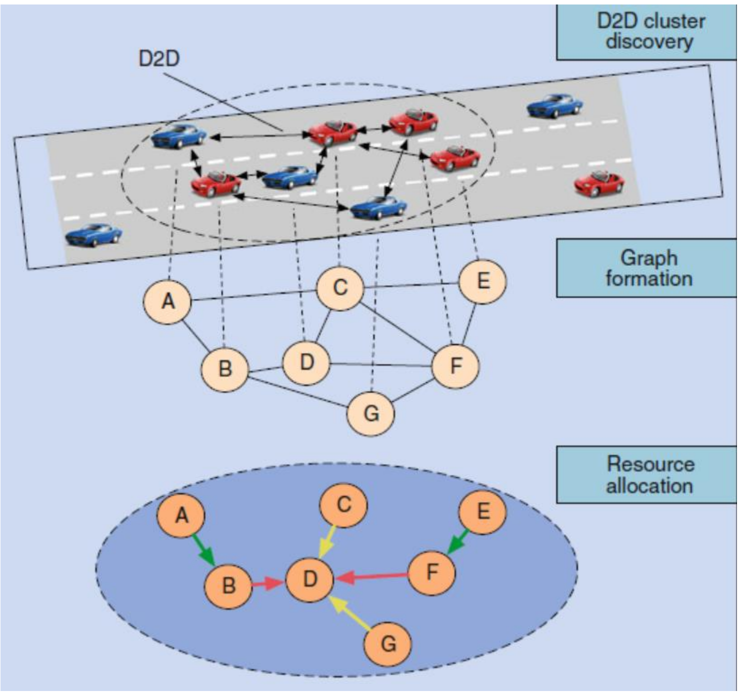


Figure 8. Proposed network management frameworks. [6]

1. D2D cluster discovery phase: In this phase eNB identifies all D2D users.

2. Graph formation phase: In this phase, the optimal routing scheme is done, based on the specific services which have to be activated.

3. Resource allocation: The resource allocation algorithm is aware of the previously established D2D network configuration.

Overall process has to be updated depending on the mobility pattern, which effects the service definition and connectivity. For the depicted in Fig. 8 scenario, eNB serves both vehicles and pedestrian users. Also all users can operate in both basic cellular mode and direct mode (MD2D). It is focused on a specific use case, in which all nodes send traffic information to a unique collector in charge of a data aggregating and transmitting to eNB. Also, because of the transmit power contrains, there is no direct communication between the sensing and collecting nodes, so a multi-hop communication scheme and an optimized routing schemes are implied. Also, resources dedicated to D2D and cellular users are mutually orthogonal.

A) D2D Cluster Discovery

Network assisted approach, via the eNB can be used to support the cluster discovery phase, since the selection of the D2D candidates is performed by the network entities. eNB decides how the clusters are formed, according to the information provided by each vehicle (i.e location, final destination, direction, speed and etc.). For example if half of the vehicles are moving toward one direction and the other half toward the opposite direction, they can be portioned into two different clusters, based on the moving directions. eNB could also keep a registry of services and resources made available by the members of the cluster in order to facilitate the providing of services within a cluster through the steps of service discovery and provider selection. Once the users are combined into a cluster (the cluster is discovered), an authentication procedure is applied in order to allow direct communication between devices in the LTE-A network. A direct beaconing is necessary in order to measure and disseminate the channel state information (CSI) for all the potential links between the cluster members. eNB is also assisting when a new vehicle is joining an already formed cluster. When a vehicle is approach an eNB, it's taking into account the context information delivered by the vehicle and can add it in the suitable cluster.

B) Graph Construction

When the eNB collects the information about the cooperating nodes, the network is selected according to the specific use case. After that this topology is optimized based on a particular target function. In vehicular social networks (VSNs) all the possible topologies are inherently data centric, because the information affects the logical connectivity and the message passing scheme. For example, data aggregation can be effectively performed through a tree topology, where a single node is collecting and refining data sensed by other one. Also, a Routing Protocol for Low-power and Lossy networks (RPL) is adopted in order to establish a loop free route toward a given destination.

C) Resource Allocation

When the topology is established, potential bottlenecks represent a serious drawback. For three topologies it involves the nodes with lower rank values (i.e nodes closer to the root (coolector)). In order to overcome this, a Network-aware resource allocation algorithm should be adopted. It should take into account the rank value of each node and provide more resources to the nodes closer to the collector. The resources are assigned according to proportional fair scheduling, where the scheduling weight w_i for the i-th node is calculated by the following formula:

$$w_i = \frac{r_i + \sum_{j \epsilon J_i} r_j}{\sum_{k \epsilon J} \; r_k + \sum_{j \epsilon J_k} r_j} \qquad (1)$$

r_i- estimated rate of the i-th node

J_i- estimated rate of the sub-tree rooted at the i-th node (J represent the overall tree)

2.3.4. Performance analysis

In [6] the performance of the proposed approach is investigated. The investigation is focused on a low-to-moderate mobility scenario, where for VANET traffic only one RB is reserved, with a minimum power transmission and a robust MCS. It refers to an anycast data dissemination, where several topologies are possible.

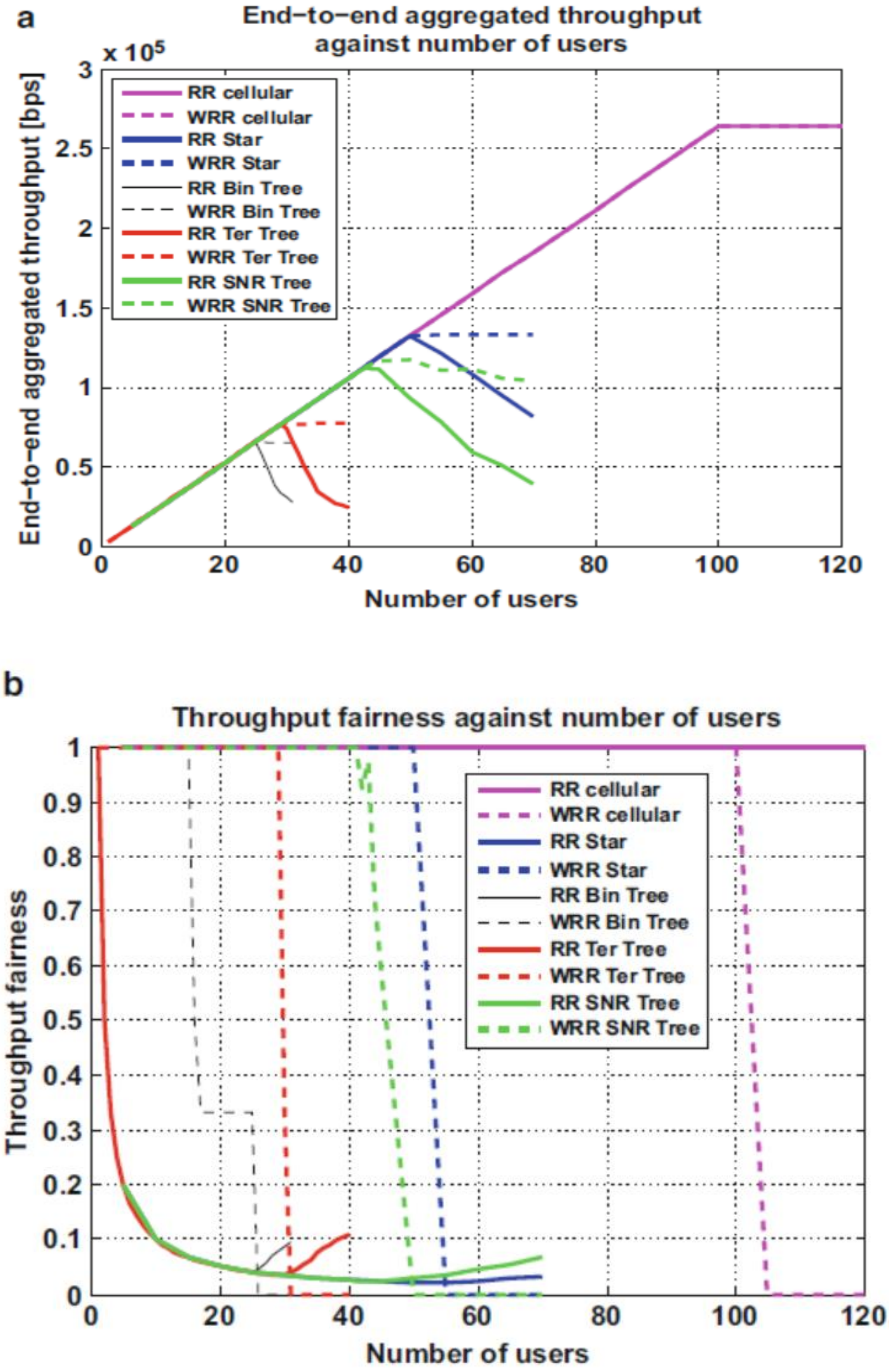


Figure 9. End-to-end aggregated throughput (a) and fairness comparisons (b) for a practical case study[6]

Fig. 9 shows the performance in terms of throughput and fairness, investigated for different network topologies. Also network unaware (RR) and aware (WRR) scheduling approaches are compared. It can be noticed that the throughput is increasing for all topologies, but the WRR scheduling scheme achieves better performance than RR. It should be mentioned that the availability of one RB allows effective management of maximum 100 vehicles that are arranged in a star topology, while the cluster size decrease for tree topology.

The performance is also investigate for a different scenario, where two traffic flows are considered. A background data and a sporadic alarm flows, with different priorities and rates.

Fig. 10 shows the investigation of performance in the case of a star and binary tree topologies with and without alarm flow for WRR and HWRR algorithms, but with unchanged network capacity. .

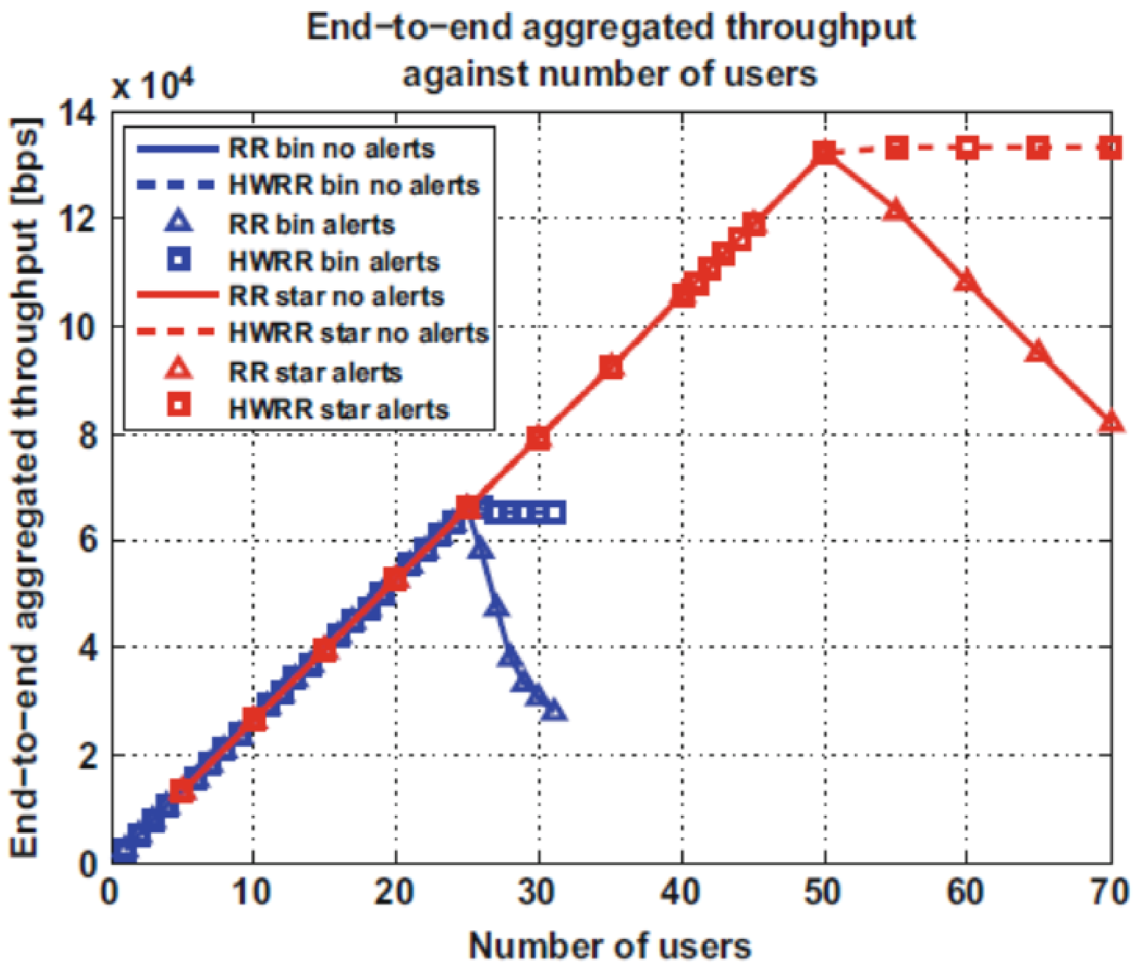
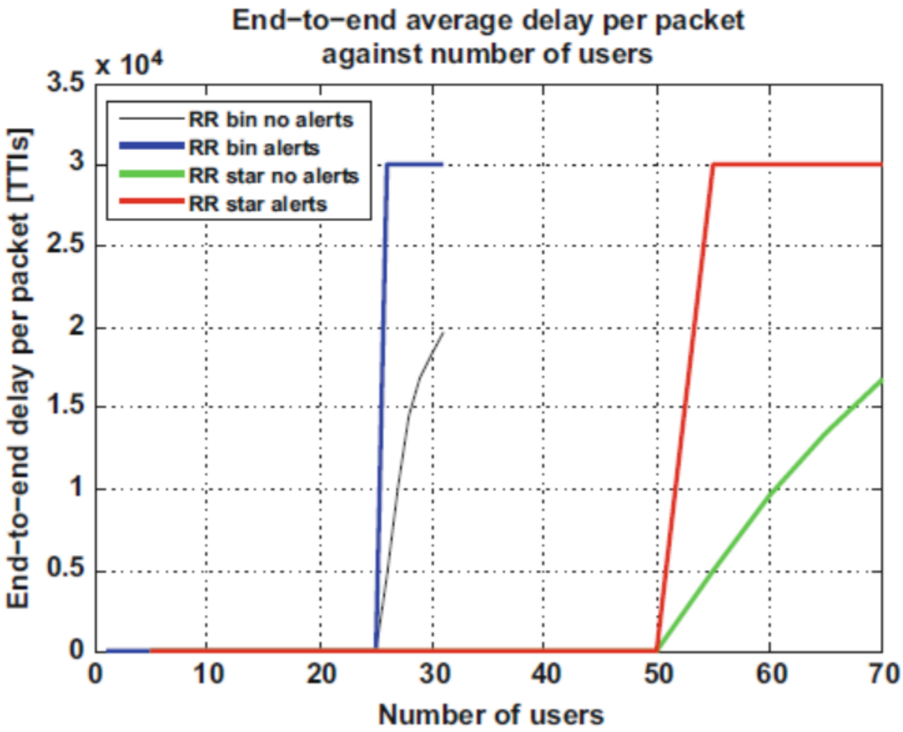


Figure 10. Throughput performance in the case of a star and a binary tree topologies with and without alarm flow for WRR and HWRR algorithms [6]

Delivery delay performance is investigated in Fig. 11. It is also evaluated for WRR and HWRR algorithms. We can see that the HWRR approach is more capable of handling flor priority and matching time constraints.



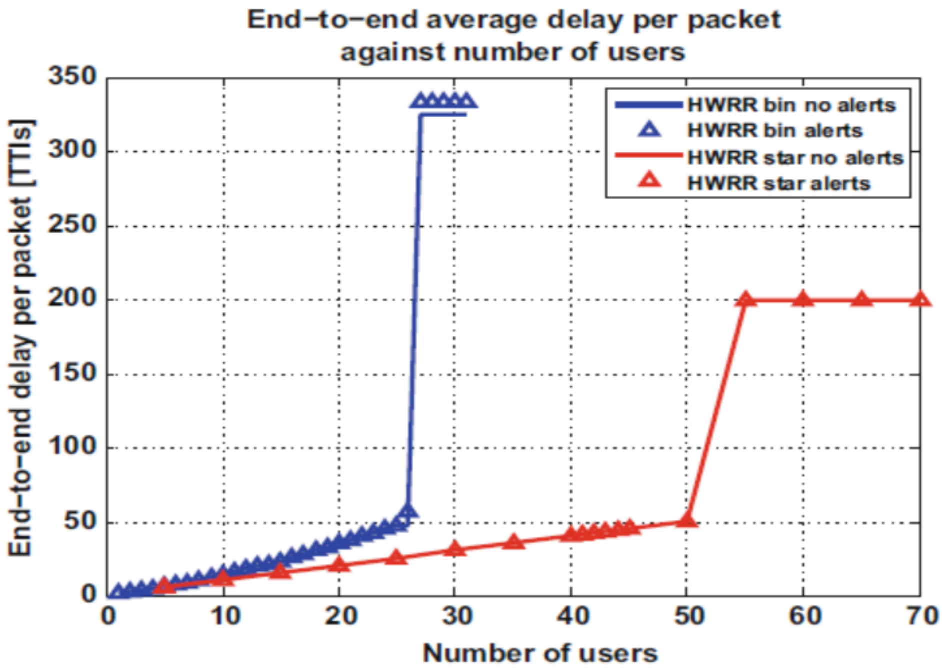


Figure 11. End-to-end delay per packet performance in the case of a star and a binary tree topologies with and without alarm flow for WRR and HWRR algorithms [6].

2.4. Advanced Communications Protocol Design

2.4.1. Small-World Networking

D2D communications has been investigated in many researches, which are focused mainly on downlink issues. However, these researches are mainly based on single-hop D2D communications, thus the development of protocols which are able to support information sharing in D2D communication is still an open issue. Some researches on large-scale self-organized distributed networks, have proposed an alternative perspective for routing protocol designs. It implies the so called small-world network structure, based on the typical for VANET deployments scale-free properties.

The analysis of the network clustering coefficient, measures the quantity of a specific group of devices and motivates the introductions of clustering based protocols. A joint clustering and routing scheme can dynamically organizes the nodes into groups (clusters). Each cluster is coordinated by one of the vehicles and is called cluster head (CH), and the rest of the vehicles are called ordinary nodes (ONs). The CH selection is based on several parameters, such as position, direction, speed, ID, energy level and etc. Several approaches to CH selections has been proposed (i.e. lowest-ID, node-weight heuristics, highest-degree and etc), but they are not always optimal solution, so the approach has to be chosen separately for each case.

2.4.2. Game Theory Inspired Approaches

For modeling the interaction among nodes in autonomic networks, Game theory (GT) is a very promising approach. In GT approaches the interaction among individual rational decision makers (players) is addressed. GT models can be divided on 2 types (Non-cooperative and Cooperative models).

In non-cooperative models players select the strategy which will maximize their own performance without taking in an account the global network performance. It leads to a steady-state equilibrium, which is not socially desirable.

In Cooperative GT, players act like a single entity. In [8] a novel classification of coalitional games is proposed. It’s main class is that of the canonical games, where the formation of large coalitions is not detrimental to any of the players. The main idea of a canonical game is to study the fairness and stability of the grand coalitions (for example in a collation of all players the stability will be achieved by finding a payoff allocation which guarantees that no player will leave the grand coalition).

GT and Coalitional games can be successfully applied for D2D communications. They can be used for solving the problems of uplink radio resource allocation, when D2D and cellular users share the available resource [9] or by achieving energy-efficient D2D communication in LTE public safety networks [10].

In order to overcome the limitations of centralized optimization and heuristic solutions for clustering schemes, so called coalitional graph games has been introduced. In this approach a specific graph interconnects the players in order to maximize their individual payoffs. It has been applied in several wireless-oriented applications (i.e for improving the physical layer security, power allocation and decentralized joint relay). However, this type of approaches have some disadvantages, such as requiring the knowledge about the actions of the other players. The solution is evaluated by each player’s point of view, which does not apply for the two different sets of players (CHs and ONs) in the VANET clustering scheme.

However these limitations can be overcome by implementing a matching theory technique into game theory and optimization [12,13].

2.4.2.1. Example of Game Theory Algorithm for VANET

In reference [13] a game theory based algorithm for VANET is proposed. The used scenario is similar to the VANET clustering scheme with two different sets of players. In it, system level and game theory modeling for both Relay Vehicles (RVs) and Source Vehicles (SVs) are described. In order to do the data forwarding, the SV has to select an optimal RV, but due to the characteristics of VANET (i.e high-speed mobility, various QoS and channel fading) there are some difficulties on the RVs selection. So they have proposed a game theory algorithm in order to overcome this difficulties.

A) System Level

The VANET in the used scenario consists of multiple SVs and RVs, but one acces point (AP), as shown in Fig. 12.

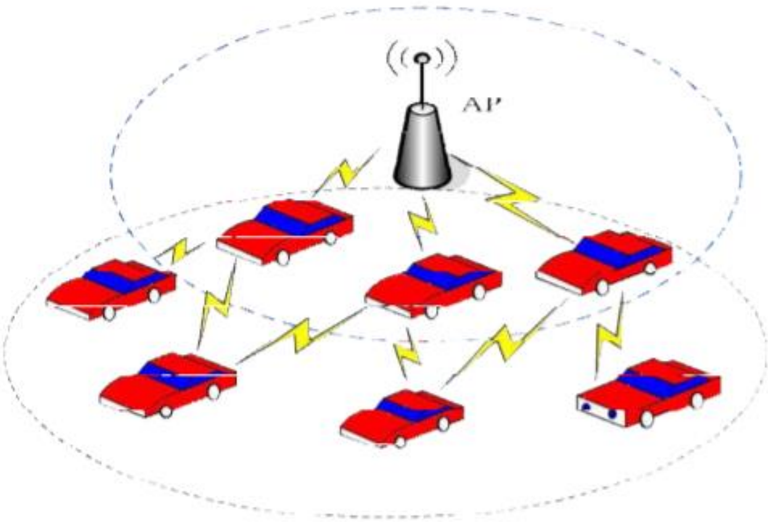


Figure 12. VANET Network model [11]

The number of RVs is N and the total number of SVs is M. Also, only there is only two-hops connection between the SVs and AP.

Assuming the number of RVs is N and the total number of SVs is M. For convenience, in this paper, only two-hop connections from the M SVs to the AP is considered, the proposed RV selection scheme can be extended to multi-hop cases.

B) Game Theory-Based Modeling for Relay Selection

The game model based on the system model from above can be formulated as follows:

- Game players are all SVs (M) and RVs (N)
- The strategy chosen from each player is as follows: on SV chooses a RV to which to forward the data and the RV can either accept or reject the request from the SV. In order to simplify the model, one

SV can access only one RV for data forwarding and the RV can forward data for only one SV.

So if the i-th SV (1≤ i ≤ M) chooses the j-th RV (1≤ j ≤ N), the payoff function of the mentioned SV and RV are U_{ij}^{SV} and U_{ij}^{RV}.

C) Payoff function modeling

The payoff function of the i-th SV which is choosing the j-th RV is:

$$U_{ij}^{SV} = W_{ij}^{S} - C_{ij}^{S},\ 1\le i \le M,\ 1\le j \le N \tag{2}$$

W_{ij}^{S} is the revenue which the i-th SV receives when applying RV forwarding scheme, and can be expressed as:

$$W_{ij}^{S} = \begin{cases} \alpha_i F_{ij}\,(1 - P_{ij}^{C})R_{i-j-AP}, \mathrm{RET}_{ij} \ge T_{ij} \\ \dfrac{\alpha_i F_{ij}\,(1 - P_{ij}^{C})R_{i-j-AP}\mathrm{RET}_{ij}}{T_{ij}}, \quad \mathrm{RET}_{ij} < T_{ij} \end{cases} \tag{3}$$

where a_i is the unit price of the i-th SV, F_{ij} is the bandwidth characteristics of the j-th RV. QoS cannot be guaranteed when the available bandwidth of the RV is smaller than the minimum bandwidth needed by the SV, so $F_{ij}=0$. If the available RV bandwidth is bigger than the maximum required by the SV bandwidth, the ideal provision can be expected and $F_{ij}=1$.

When the available RV bandwidth falls between the two cases explained above, F_{ij} is described by the following expression:

$$F_{ij} = \frac{\left(B_j^R - B_i^{min}\right)}{(B_i^{max} - B_i^{min})},\ B_i^{min} < B_j^R < B_i^{max} \tag{4}$$

In formula 2, T_{ij} is the time required for i-th SV to forward packets to j-th RV. It can be expressed by the following formula:

$$T_{ij} = 2T_{OH} + \frac{S_{PL}+S_{MAC}}{R_{SR}} + \frac{S_{PL}+S_{MAC}}{R_{RA}} + T_{\mathrm{BO}} + T_{SIFS}, \tag{5}$$

where $T_{OH} = T_{DIFS} + 3T_{SIFS} + T_{RTS} + T_{CTS} + T_{ACK} + 2T_{PLCP}$. (6)

-T_{SIFS} and T_{DIFS} are short inter-frame space and distributed inter frame space which arew defined in IEEE 802.11p.

-T_{PLCP} is the duration of the physical layer convergence procedure

-T_{RTS} and T_{CTS} are the duration of the RTS and CTS frames.

-T_{ACK} is the duration of one ACK frame.

In (5) S_{PL} and S_{MAC} represent the size of an MAC payload and header, respectively. R_{SR} and R_{RA} represent the the data transmission link between SV and RV, and T_{BO} represents the average backoff time. R_{i-j-AP} represents the achievable transmission rate between the i-th SV and AP through the j-th RV, and is explained by the following expression:

$$R_{i-j-AP} = \min(R_{i-j}, R_{j-AP}). \tag{7}$$

R_{i-j} *and* R_{j-AP} represent the achievable rate of the link between i-th SV and j-th RV and between the j-th RV and the AP. $R_{i\text{-}j}$ and $R_{j\text{-}AP}$ are calculated according to the Shannon-Harley Theorem:

$$R_{i-j} = B_j^R log_2(1 + SNR_{i-j}) \tag{8}$$

$$R_{j-AP} = B_j^R log_2\left(1 + SNR_{j-AP}\right)$$

where B represents the bandwidth (in Hz) of the j-th RV and SNR represents the signal noise ratio of the links.

In (2) C_{ij}^S, represents the service fee that the SV needs to afford, in order to accomplish the data-forwarding. It can be calculated by the following formula:

$$C_{ij}^S = \beta_j F_{ij}^k \tag{9}$$

where β_j the bandwidth price factor for forwarding packets and k is the index of the bandwidth effect.

The payoff function of the j-th RV which is offering data forwarding for SVs is:

$$U_{ij}^{RV} = W_{ij}^R - C_{ij}^R,\ 1 \leq \mathrm{i} \leq \mathrm{M},\ 1 \leq \mathrm{j} \leq \mathrm{N} \tag{10}$$

The revenue that j-th RV receives from the data forwarding is represented by the following expression:

$$W_{ij}^R = P_{j-AP}^S P_{i-j}^S \beta_j F_{i-j} \tag{11}$$

where P_{j-AP}^S *and* P_{i-j}^S represents the probability for a successful transmission between the j-th RV and the AP, and between the i-th SV and j-th RV. The probabilities are calculate by the following formulas:

$$P_{jAP}^S = \Pr(\frac{E_s}{N_0} h_{jAP}^2 > \psi_{th1}) = 1 - \frac{\gamma(m, \frac{mN_0 d^\theta \psi_{th1}}{cE_s})}{\Gamma(m)} \tag{12}$$

$$P_{ij}^S = \Pr(\frac{E_s}{N_0} h_{ij}{}^2 > \psi_{th2}) = 1 - \frac{G_{1,3}^{2,1}[\frac{m_1 m_2 N_0 \psi_{th2}}{\Omega_1 \Omega_2 E_s}\Big|_{m_1, m_2, 0}^{1}]}{\Gamma(m_1)\Gamma(m_2)}$$

In (10) C_{ij}^R represents the forwarding cost of the j-th RV which is forwarding the data from the i-th SV. The nonlinear relationship between the undertaken cost and the factors that are affecting the cost can be evaluated through the Sigmoid function, and the cost can be calculated through the following expression:

$$C_{ij}^R = \frac{c_j F_{i-j}}{1 + e^{\theta_j(\varphi_j - P_{j-AP}^S P_{i-j}^S B_i^{max})}} \tag{13}$$

where θ_j and φ_j determine the steepness and the inflection point of the RV cost curve, while c_j is the cost factor of the j-th RV.

3.2. Simulator Structure

3.2.1. A Typical Simulation

Individual parts of the simulator are defined by classes. This allows adding new functions without the need of altering the other functions. Overview of the simulator’s structure is shown on Fig.7.

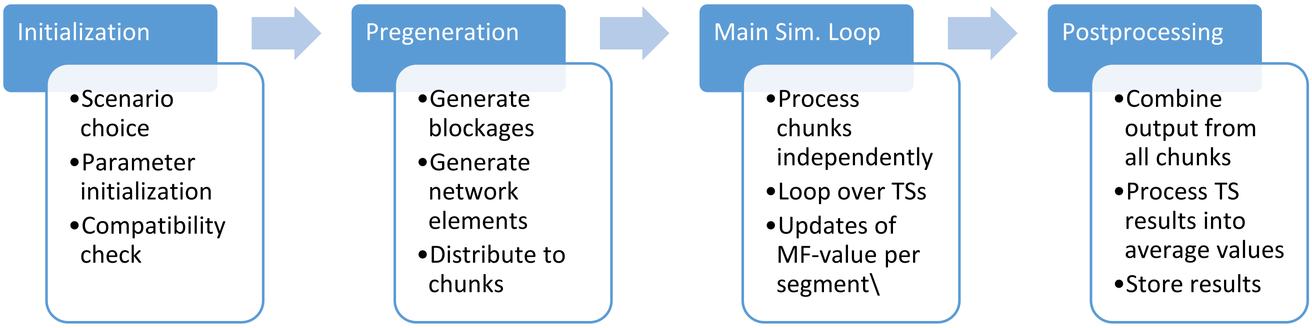


Figure 13. Overview of the main parts of the simulator [12]

The simulation starts by executing the desired simulation launcher file. It already fixes the parameter set for the simulation. Some of the parameters are defined in the scenario file while others are predefined in the classes. In the file simulate.m the simulation object is creadted and the predefined parameters are attached to this object.

```
% create simulation object
localSimulation = simulation . LocalSimulation ( params );
```

After that the time line and the network element objects of the simulation are generated. This is done in:

```
% setup simulation and generate network elements
localSimulation . setup ();
```

If necessary, the fast fading traces are pregenerated too.

After the pregeneration is done, the simulation is carried out. The main simulation loop (over chunks and time slots) is generated by the following lines:

```
% main simulation loop
localSimulation .run ();
```

Random samples in time and/or in space are simulated and the results for each time slot are stored. The time samples are represented by the channel realizations and scheduling decisions, while the space samples are represented by the geometry of the network elements. After the main loop simulation comes the post processing. It is executed by the following lines of code:

```
% main simulation loop
localSimulation .run ();
```

In the post processing the individual results from main the main simulation loop are stored in the “results” object. Only the predefined in the simulation parameters results will be extracted, calculated and stored. After that the results we choose are plotted. The code lines for plotting should be in the simulation launcher.

3.2.2.The simulator time line.

The time line of the simulator is divided into 3 different units: chunks, segments and timeslots (TSs).

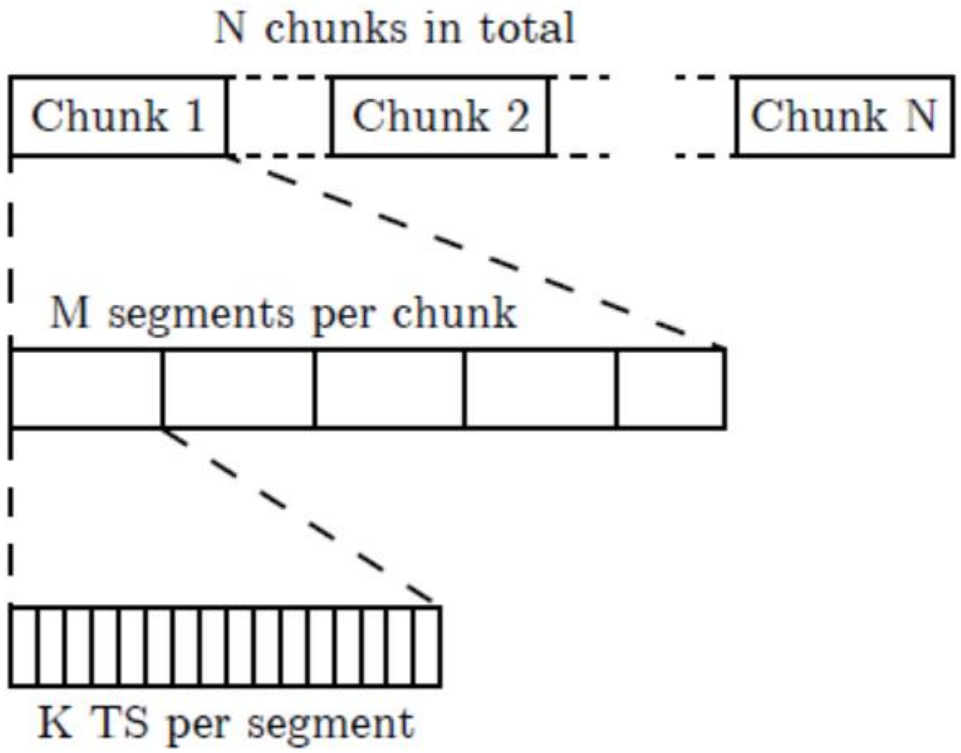


Figure 14. Simulator time units [12]

The TS is the shortest one and correspondents to the scheduling granularity. Thus it equals to one iteration of the inner simulation loop. By default the length of one TS is 1ms in order to represent an LTE-A subfame, but it can be specified freely if necessary. The segment is consisted of various TSs and corresponds to the time and distance in which the macroscopic fading (MF) is assumed to be constant (e.g. the user association, large scale path loss value and etc.). MF values are updated at the beginning of each segment. The length of the each segment depends of the user trajectory, speed and the correlation distance between the MF values.

Each chunk is consisted of fixed number of time slots and 1 or more segments (depending on the maximum user speed). For creating the user trejactory, a consecutive generation is assumed, the distance between chunk is considered too. This leads to an uncorrelated user positions among chunks. Even for the stationary scenarios, the channel coefficient are not correlated, due to the non-constant scattering environment. The chunks are also the basis for the parallel simulations, since all of the necessary data is independent for each chunk and the results are calculated without taking into account the results in the previous chunks.

The time line is utilized by the following lines of code:

```
params . time . numberOfChunks = 2; % number of chunks in the simulation
params . time . slotDuration = 1e -3; % time slots duration in seconds
params . time . slotsPerChunk = 10; % each chunk consists of that many slots
params . time . timeBetweenChunksInSlots = 50; % timespan between two simulated time chunks
```

In the simulation.ChunkSimulation.m, the following function is contained:

```
function setNewSegmentIndicator (obj)
```

It creates a logical array that is true for all first in a segment slots. Also it Marks all slots, for which the user has moved further than the maximum correlation distance. This means that based on the trajectory and correlation distance an indicator is set, where a new segment starts.

3.2.3. Main Simulation Loop

The main simulation loop contains a loop over chunks, whereas each chunk contains a loop over TSs (cf. Fig. 9).

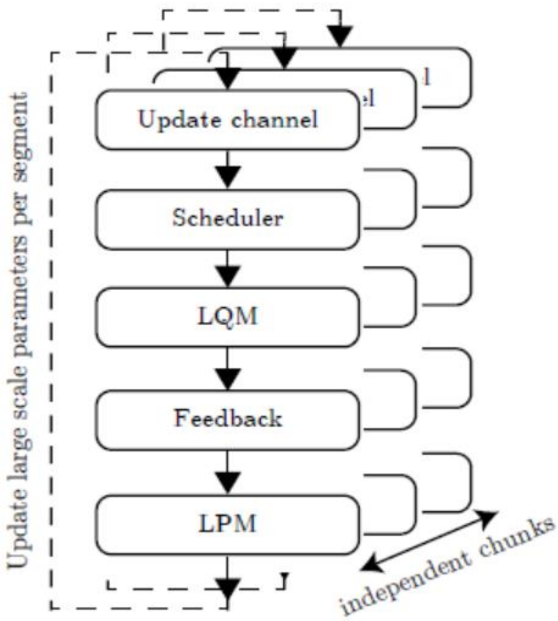


Figure 15. Main loop simulation [12]

The function simulation.LocalSimulation.run defines the loop over chunks:

```
% initialitze chunk result list (for local simulation)
chunkResultList (obj. parameters . time . numberOfChunks ) = simulation . ChunkResult ;
% run simulation for each chunk
for ii = 1: obj . parameters . time . numberOfChunks
fprintf (' simulating chunk %d ...\ n', ii);
obj . chunkSimulationList = [obj. chunkSimulationList, simulation . ChunkSimulation (obj. simulationSetup.chunkConfigList (ii))];
chunkResultList (ii) = obj. chunkSimulationList (ii);
runSimulation ();
end
```

In cases of parallel simulation, the loop over chunks is defined by

```
% prepare chunk simulation list
chunkSimulationList = [];
for ii = 1:( obj. parameters . time . numberOfChunks )
chunkSimulationList = [ chunkSimulationList , simulation .ChunkSimulation (obj. simulationSetup . chunkConfigList (ii))];
end
% run simulations for all chunks
parfor ii = 1: obj . parameters . time . numberOfChunks
fprintf (' simulating chunk %d ...\ n', ii);
chunkResultList (ii) = chunkSimulationList (ii). runSimulation ();
end
```

simulation.ParallelSimulation.run:

For both local and parallel simulation, the data is distributed among chunks and stored in chunkSimulationList.
Regardless of the type of simulation, each chunk contains the inner loop over the time slots.
This loop is defined in simulation.ChunkSimulation.runSimulation:

```
% main simulation loop
for iSlot = 1: obj. NSlots
% In a new segment the cell association can change and
% handovers need to be performed .
if obj. chunkConfig . isNewSegment ( iSlot )
% update cell association if this is new segment
obj . updateUsersAttachedToBaseStations ( iSlot );
% find cells that belong in the interference region
obj . filterPureInterferenceCells ( iSlot );
% perform handovers - clear feedback buffers of users in new cell
obj . performHandover ( iSlot );
end % if this slot is the first one in a segment
```

Another segment update is performed per user and sets the appropriate (Macroscopic fading) MF values for each user for the current segment:

```
if obj. chunkConfig . isNewSegment ( iSlot )
  bb = obj. userToBSassignmentArrayDL (iUE , obj . getiSegment ( iSlot));
  desired = false (1, obj . nAntennas );
  desired (bb) = true ; % only works for one antenna per bs
  % get new macroscopic parameters
  thisShadowFadingdB = obj. shadowFadingdB (:, iUE , iSlot ).';
  pathlossdB = obj. pathLossTableDL (:, iUE , obj. getiSegment (iSlot )).' + obj. wallLossdB (:, iUE , obj. getiSegment ( iSlot )).';
  gaindB = obj . antennaGaindB (:, iUE , obj . getiSegment ( iSlot )).';
  % update macroscopic parameters
  LQMDL (iUE). updateMacroscopic ( desired , pathlossdB , gaindB ,thisShadowFadingdB );
end
```

3.3. Key functionalities

3.3.1. Generation of network elements and geometry

In order to evaluate the performance of a large network consisted of a substantial number of BSs and users, at its core, the Vienna 5G SLS simulates the communication between base stations and users. The coexistence of different types of users and BSs is possible, which allows the simulation of multi-tier networks, and support of more diverse and realistic scenarios. Additionally to the various propagation models, there is option to distribute blockage objects and use the geometry to calculate different propagation parameters.

A) Base stations

In the simulation the Bs and antenna are different objects. Each BS can have one or more attached antennas, which can be seen as physical entities with a position f(x,y,z) in the 3D space. The antenna object is consisted of an antenna array with a number of transmit (N_{Tx}) and receive antennas (N_{Rx}). The physical position of the BS is replaced by the physical positions specified on the assigned antenna objects. Also the structures enables the simulation of Distributed Antenna Systems (DAS) and Remote Radio Heads (RRH) without any additional extensions.

The BS object is defined in +networkElements/+bs/BaseStation.m and has the following characterizing properties:

-antennaList - list of all attached antennas

-attachedUsers- a list of users that are connected to this BS and all of their properties

-type- the type of the BS (i.e pico, femto, macro)

-isRoi – indicator for BS that is in the Region of Interests (ROI)

+networkGeometry defines all of the placement methods that can be used. The choice of placement method, should be specified in the scenario file. Each BS type is indicated as an integer in +parameters/+setting/BaseStationType.m. The default number of available options are 3 (1. Macro, 2. Pico, 3. Femto), when a new type of BS is added, it has also to be indicated with a corresponding enumeration.

+networkElements/+bs/Antenna.m defines the antenna object. Each antenna has the following properties:

-id- integer identifier of the antenna

-baseStationType - type of the BS to which the antenna is attached

-usedCCs - carriers on which this BS transmits

-nTX - number of transmit antennas

-nRX - number of receive antennas

-alwaysOn – This is a Boolean value that indicates if BS is always transmitting. If this value is true (1) the attached antennas are always transmitting, respectively are always generating interference, even if no users are attached to the BS in the current time slot.

-rbGrid – provides the scheduling information and details about the resource blocks

-gainBmax – maximum antenna gain in dBi

- azimuth- the angle in degrees in which antenna has its maximum gain

B) Users

Users and other endpoints of the communication link are defined in +networkElements/+ue/User.m. Each user has the following properties:

-id - integer identification number

-nRX – number of receive antennas of the user

-nTX –number of transmit antennas of the user

-txMode - transmission mode

-transmitPower – transmit power

-channelModelTypes – Indicates the possible channel models that can be used by the user. All supported channel models are defined in parameters.setting.ChannelModel

-currentChannelModel – channel model used by the user in the current time slot

-speed – user speed (m/s)

Similar to BSs, different placement methods can be used for users too. The placement methods are defined in +networkGeometry, and should be specified in the scenario file.

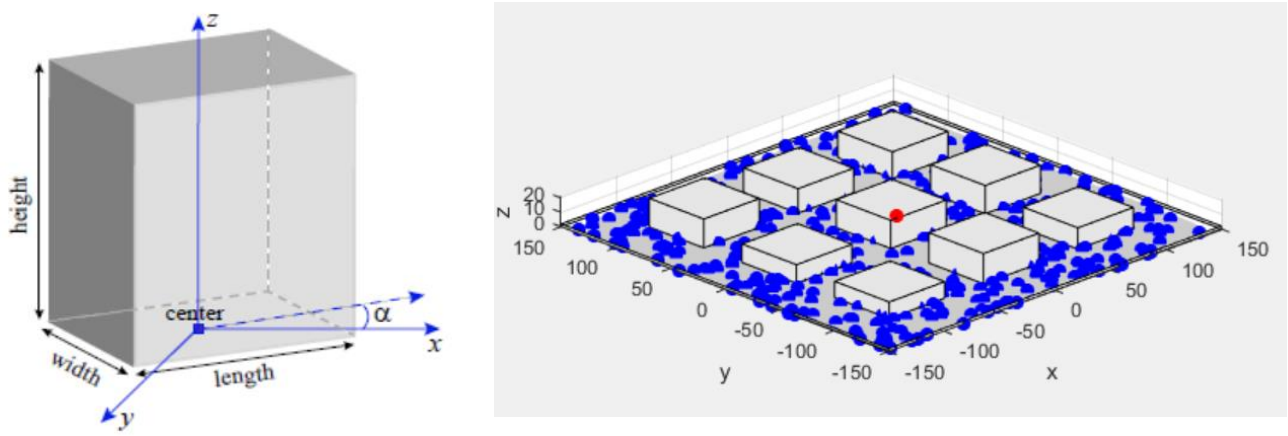
C) Blockages

Additionally to the network element generation, the Vienna 5G SLS supports the generation of blockages. Blockages superclass is defined in +blockages/Blockage.m. The basic signal blocking object is a rectangular wall with random dimensions and positions in 3D. Building are created by a combination of 5 walls (4 walls on each side and a ceiling). Thanks to these basic elements, a city layout such as Manhattan grid with streets and building blocks can be generated.

Building objects are defined in +blockages/RectangularBudiling.m and each building has the following properties:

-length – x-axis
-width – y-axis
-height- z-axis
-angle- angle in radians according to which the building is rotated with respect to the ground plane
- center- position (x;y) of the building center

Additionally, buildings can be generated together with streets in order to simulate an Urban scenario. Example for that is the ManhattanGrid scenario, where the building and sreets are generated according to Manhattan grid layout (Fig 10b).



A) Rectangular building blockages *B) Manhattan grid scenario*

Figure 16. Blockages [12]

3.3.2. Link Quality and Link Performance Model

A) Link Quality Model

The Link Quality Model (LQM) quantifies the quality of a link in two SINR Measures (post equalization SINR and wideband SINR):

-The post equalization SINR is calculated for each layer of each resource block. It takes in account the utilized receive filters and precoder, the inter layer interference and the transmit power allocation.

- The wideband SINR represents the quality of all physical resources in one timeslot. It is calculated based on large scale parameters and takes in account the antenna gain, path loss, noise and transmit power.

The output of LQM is shown Fig.11.

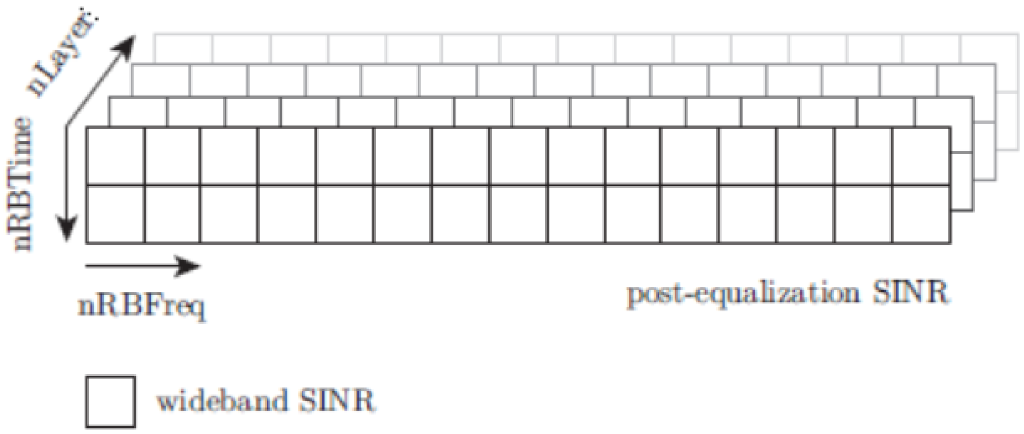


Figure 17. The dimensions of the output of the LQM [12]

To calculate those SINR values, LQM needs constant parameters (i.e. receiver noise figure) and large scale parameters (i.e. path loss and antenna gain)a nd small scale parameters (i.e. channel matrix and transmit power allocation). The constant parameters are set in the class constructor and they are constants during the whole simulation, while the small and large scale parameters have to be updated. Large scale parameters are updated for each segment with the function updateMacroscopic and small scale channel parameters are updated for each time slot with function updateSmallScale. A simplified scheme of the functionality of LQM is depicted in Fig. 12.

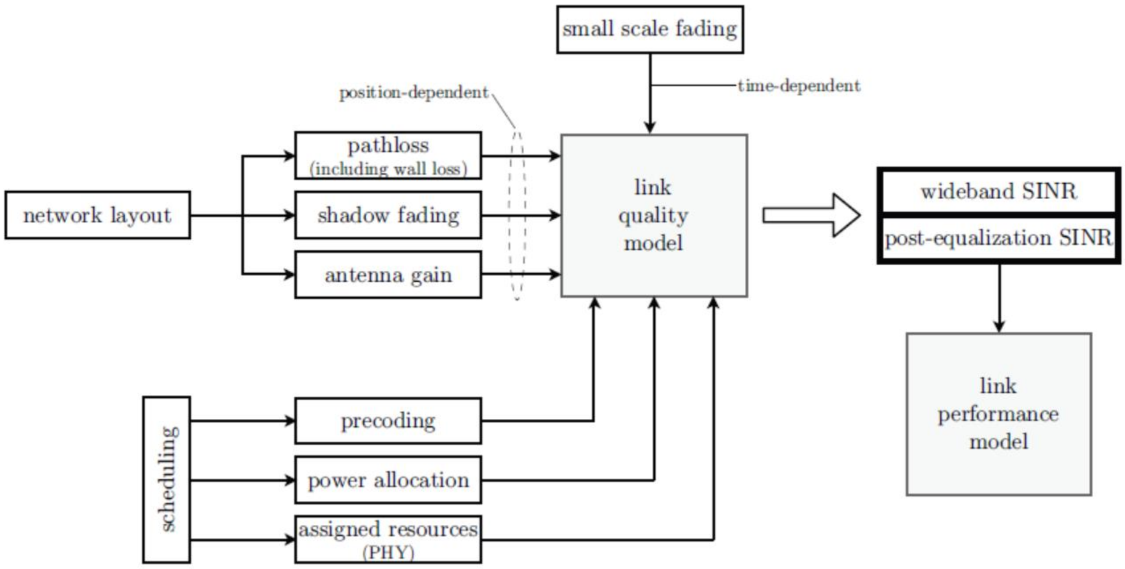


Figure 18. Schematic of Link Quality Model [12]

For the implementation in Fig.12 a Zero Forcing receiver is used and the small scale fading is a constant for the time duration of a timeslot. It is also assumed that if no user is scheduled for an interfering BS, the BS does not generate any interference (unless the alwaysOn property is set as 1). Also one receiver operates on a constant number of layers within a timeslot and the allocated transmit power is equally distributed between these layers.

B) Link Performance Model

The Link Performance Model determines the throughput of a given user based on the SINR values calculated by LQM and the decision of the scheduler. This process is done in three steps:

1. Calculate an average SINR value for each transmitted codeword over the scheduled RBs and transmission layers.

2. Map the average SINR to a Block Error Ratio (BLER) value

3. Determines the number of transmitted bits based on the Channel Quality Indicator and BLER.

The third is based on a Bernoulli experiment that defines if the transmission was successful or not. If the transmission is successful the number of transmitted bits is set to the maximum size of the Transport Block for the scheduled RBs, otherwise it is set to zero. Depending on the chosen settings, LPM can perform Bernoulli experiment in every timeslot or give the expected value of the experiment as an output.

Besides the actual throughput the LPM calculates an upper bound on the throughput. This is the throughput that would result if the scheduler decides for the CQI that would result in the highest throughput. In order to calculate this value, LPM performs the same steps for all possible CQI values and takes the maximum over the resulting throughput.

3.3.3. Simulation of Propagation Effects

A) Path Loss Modeling and Situation dependent Model Choice

In TU-Wien 5G SLS the path loss modeling implementation assures modularity and flexibility on one hand, but a simple straight forward usage on the other. The superclass PathlossModel allows all path models to be used in the same way with the functions set in it.

The necessary parameters and available path loss models are described in the enumeration file parameters.seetings.PathlossModel. The path loss is calculated without the wall loss and antenna gain and it's independent for each chunk.

Path loss values are saved in simulation.ChunkSimulation.calculatePathloss. The function is called in the initialization phase of each chunk, right before the loop over TSs. First, the link condition (i.e. LOS or NLOS links) is determined. After that the appropriate model for the considered link type is chosen.

B) Modeling of Shadow Fading

The modeling of SF is based on Shadow Fading Maps (SFMs). Those maps are arrays of so called Shadow Fading Values (SFVs), which are spatially correlated random variables (RVs). Those variables are generated in 4 steps:

1. Generate an array of uncorrelated Gaussian RVs

2. Apply a Fast Fourier Transformation to the array

3. Multiply the transformed array with the Power Spectral Density (PSD) of the intended spatial correlation.

4. Apply an inverse Fast Fourier Transformation.

By using this method the correlated shadow fading values for positions with minimal granularity can be graduated without the need of resolving the whole region of interest (ROI) with the same precision.

C) PDP Channel Models

Used channel models are defined through a Power Delay Profile (PDP). Channel models are defined in the enumeration file parameters.setting.ChanelModelType. For the PDP channel models, channel traces are generated before the main loop simulation. The generation is done by the PDPcontainer in smallScaleFading package. In this package all number of antennas, carrier frequency and channel models that can occur during the simulation are evaluated. The resulting channel traces are saved in the folder dataFiles/channelTraces.

Channel realization for each time slot the function PDPcontainer.updatePDPcontainer loads the needed channel traces into the memory and the function PDPcontainer.getChannelForAllAntennas selects a random pard of the trac in order to realize channel for one time slot. On one hand channel traces should be longer than the total simulation time so samples taken from the trace will be independent from each other, but on other hand traces should be as shorter as possible, in order to save memory. Length of each trace is set in SmallScaleParameters class.

3.3.4. Scheduling

Basic version of scheduling is done separately for every BS, so there is always a one to one relation between an instance of the scheduler class and the BS. In the scheduler the corresponding BS is called “attached BS”.

The scheduling process in the simulator is built on top of the class Scheduler It defines the following functions:

-scheduleDL: This function allocates users and transmit power to RBs. It has to be called once for every simulation slot.

-addUsersDL: This function adds users to the scheduler queue. It has to be called when a user connects to the attached BS.

-removeUsersDL: This function removes users from the scheduler. It has to be called when a user disconnects from the attached BS.

Those functions has to be implemented for every type of scheduler. Additional to them there are two more pairs of functions that are not abstract:

-scheduleDLCommon: It’s called from within scheduled and performs all calculations that are common to all types of schedulers (i.e. decides the CQI for all RBs that are scheduled for one user).

-updateAttachedUsersDL: This function can be used instead of using addUsersDL and removeUsersDL separately. It takes a list of users that should be attached to the BS and if necessary calls the two separate functions.

3.3.5. Feedback

The Feedback class provides the scheduler with information on channel conditions (Rank Indicator(RI), CQI and Precoding Matrix Indicator (PMI)). The scheduler needs this information in order to determine optimal transmission parameters for future timeslots. All feedback types supported by the simulator are implemented as subclasses of the Feedback class. Currently there are two types of feedback:

-LTEDLFeedback: It corresponds to the feedback used in the LTE Downlink

-MinimumFeedback: This type of feedback shows the minimum of feedback required by the scheduler.

Currently, supported feedback in SLS is only for Single-Input Single-Outpu (SISO) and thus only the CQI is utilized. Each feedback type implements the method calculateFeedback. For LTE, the CQI values are computed using the reported SINR values.

The SINR to CQI mapping is implemented in subclasses of parameters.transmissionParameters.CqiParameters.

3.3.6. Post Processing

After the main simulation loop is done, by calling the function combineResults, the results for each chunk are collected and combined in the post processing phase. SLS is equipped with one postprocessor for full simulation and two postprocessors for lite simulations:

-PartialPP: This is the postprocessor for full simulation. It saves the results defined by the ResulstsPartial class plus the additional results indicated by the SaveObject parameter. The following results are saved with this postprocessor:

- Signal to Noise Ratio (SNR) and Signal to Interference plus Noise Ratio values (one value per segment)
- userThroughputMBitPerSec: contains the downlink throughput of each user in the ROI in each TS. It is calculated in Mbit/s
- userThroughputBit: downlink user throughput in bits.
- widebandSinr: the wideband SINR of all users within the ROI in each time slot in Mbit/s.
- effectiveSinr: effective downlink SINR per timeslot.
- assignedBS: assigned BS for all users per timeslot.
- Feedback: feedback parameter for all users
- schedulerSignaling: scheduling information of all users
- networkElements: contains the information about users,BSs,blockasges, buildings and street geometry
- params: contains the parameters that are defined at the beginning of the simulation
- Additional results: losMap, isIndoor,antennaBsMapper and pathlossTable

- LiteNoNetworkPP: This postprocessor is used in lite simulations when we don’t want to save the network elements. Functionality in lite simulation is reduced, so many parameters that are saved in the partialPP, are not computed. Only SNR, SINR values and the necessary for the simulation parameters are stored and saved.

-LiteWithNetworkPP: It is used also in lite simulations only. It is similar to LiteNoNetworkPP but in this case the network elements are saved too.

Default used postprocessor is PartialPP, but the postprocessor type can be changed in the scenario configuration file:

```
params . postprocessor = simulation . postprocessing .LiteNoNetworkPP ;
```

III. Implementation.

The main idea of this study is to propose a V2I traffic model with 3 types of traffic (Safety applications, Public applications and Entertainment applications). The scheduler should be able to distinguish the traffic and schedule the users based on the priority of the traffic. Safety applications' traffic is with the highest priority, while the Entertainment applications' traffic has the lowest priority (in other words it should be scheduled last). The implementation is done using TU-Wien 5G System Level Simulator and Matlab 2018b.

1. Scenario:

The simulation is based on the urban scenario specified in 3GPP TR 36.885 v14.0.0 (2016-06) standard [13]. Each 100ms the vehicles broadcast single-hop packets (the vehicles transfer packets directly to the base station) with the size of 300 bytes starting from a random initial time instance.

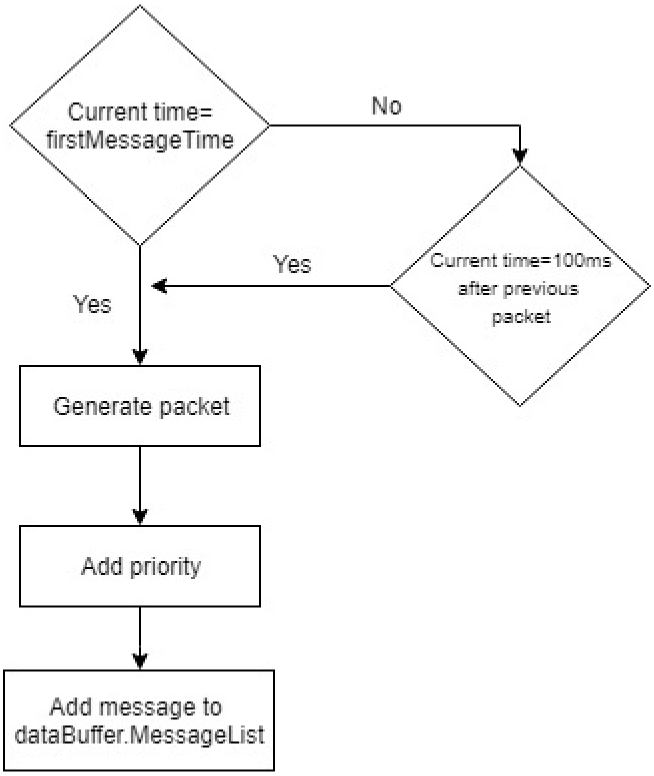


Fig. 19. Packet generator.

The simulation is done with 100, 200 and 500 users, following a Poisson arrival. Vehicles travel in a Manhattan grid road topology with a speed of 15 km/h. The path loss model in the simulation considers large scale effects, shadowing effects and buildings obstruct line-of-sight communications. The scenario parameters are based on 3GPP TR 36.885 v14.0.0 (2016-06) and are described in Table 4. All non-default for the simulator parameters are specified in ManhattanGridDataBuffersScenario.m .

Parameter	**Value**
Bandwidth	5 MHz/10 MHz/15 MHz
Message size	2400 bits
Moving speed	15 km/h
Message frequency	10 Hz
Latency requirements	100 ms
DL carrier frequency	2 GHz
UL carrier frequency	1.8 GHz

Table 4. 3GPP TR 36.885 v14.0.0 Standard parameters

Manhattan grid characteristics are described in Table 5.

Parameter	Value
Block length	50m
Block width	50m
Street width	35m
Minimum building height	10m
Maximum building height	25m
Wall Loss	10 dB
Number of users	100 / 200 / 500

Table 5. Manhattan grid parameters

2. Data buffers

Each user has an assigned data buffer, which contains information about transmitted messages. Data buffers are defined in the DataBuffer superclass and have the following properties:

-messageList: Contains the messages which have to be transmitted.

-messageSize: Specifies the size of each message in bits. By default it is 1, but for the simulation it is 2400.

-latReq: latency requirements (in ms) for each message. By default, it is 2ms, but for the simulation it is 100ms.

-sentMessages: Defines the number of successfully transmitted messages.

-discardedMessages: Defines the number of messages which are discarded due to high latency. If the message is not transmitted within 100ms (latReq), it will be discarded regardless of its priority.

-transmittedMessagesLatency: An array, which contains the latency values of the successfully transmitted messages in every timeslot.

-msgPriority: Contains the priority of last generated message.

-packetPriority: An array that contains the priority of each message generated by the user.

The user's data buffer is updated after each timeslot. If the message is completely transmitted it is removed from messageList, otherwise the data in the messageList is reduced by the number of transmitted bits.

3. Scheduler algorithm

The proposed scheduler is a modified Round Robin scheduler, with 3 types of sub-queues (safetyQueue, publicQueue, and entertainmentQueue). Each type of traffic goes in different sub-queue, after that they are combined into 1 queue where the safetyQueue is always at the beginning. The users are scheduled from the first position of the queue.

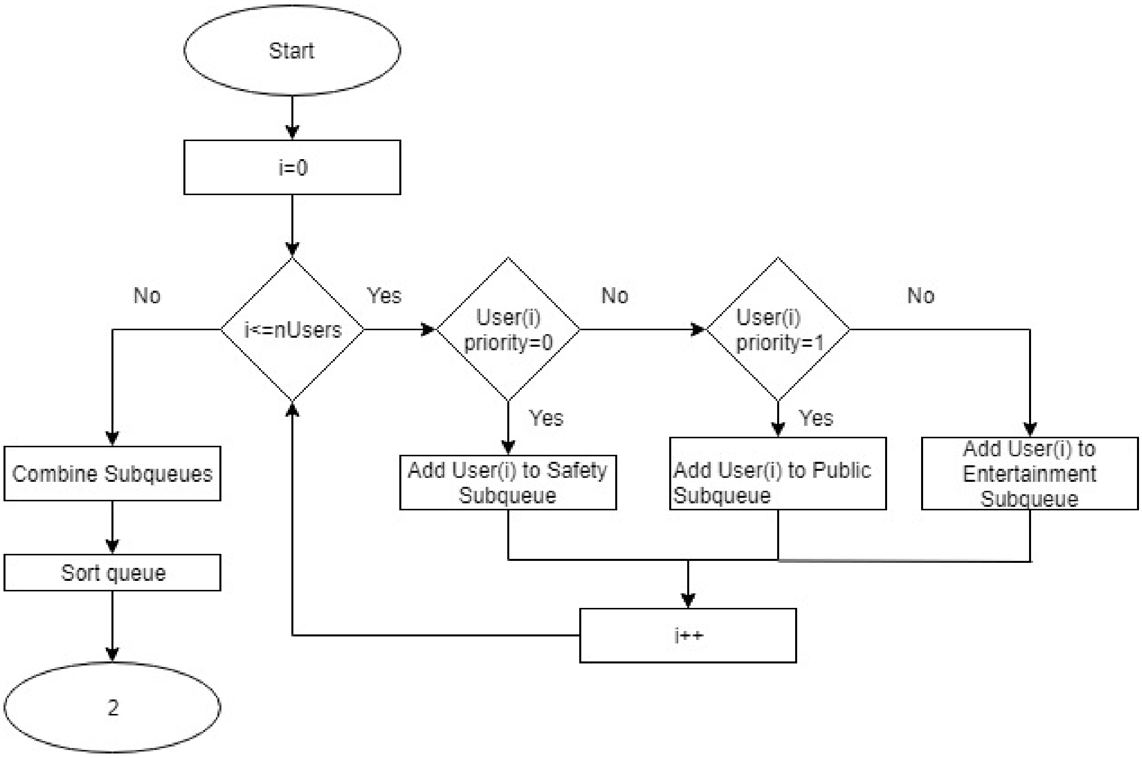


Fig 20. (1) Queuing algorithm

Users are always in the queue, as long as they are attached to the base station. However, if the messageList of the user is empty, no resources will be allocated (Fig 20). Every timeslot the scheduler is checking the messageList of each user in the queue. If the messageList is empty, no resources are allocated to this user.

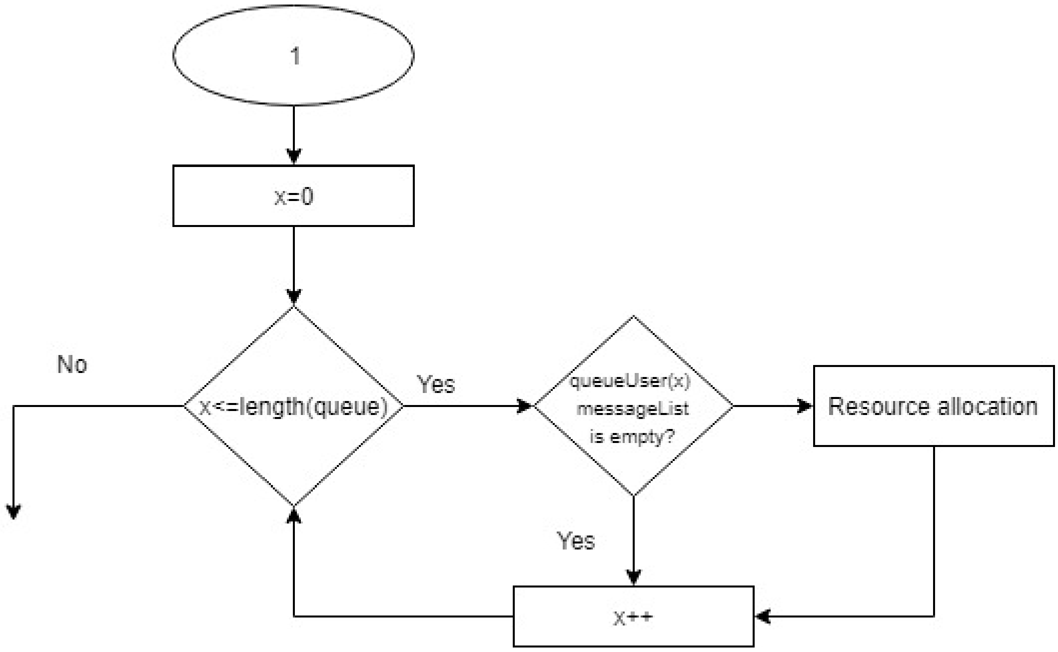


Fig 21. (2) Resource allocation according to messageList size

Notes: 1- Queuing algorithm from Fig 19.
2- Resource allocation algorithm from Fig 20.

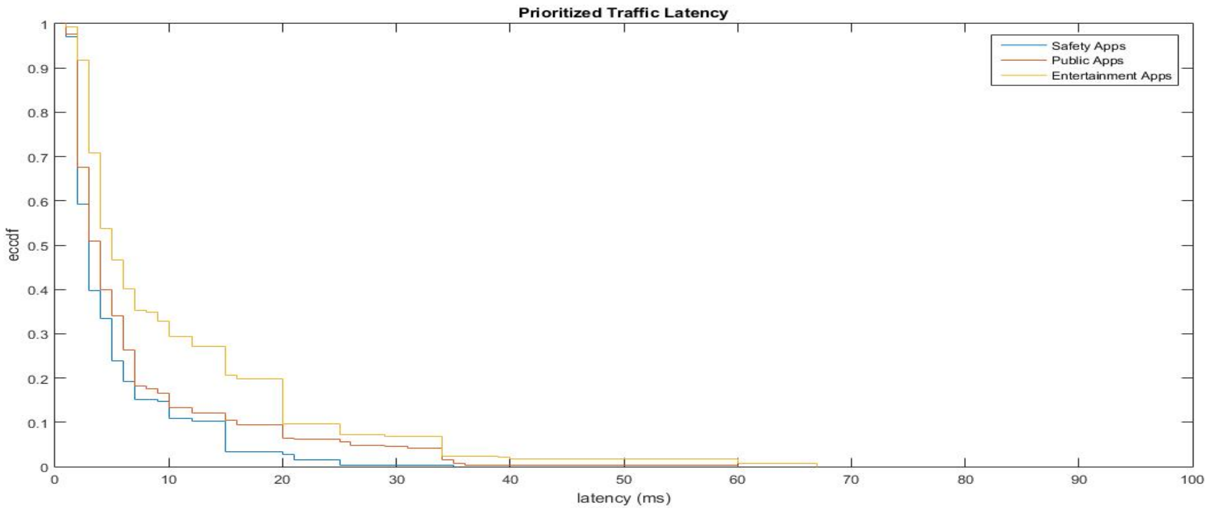
IV. Result analysis

The simulations are done with 300 timeslots and 100, 200 and 500 users. The latency results are plotted according to their Empirical Complementary Cumulative Distribution Function (ECCDF). In order to evaluate the performance of the new proposed scheduler, the results after using “priorityScheduler.m” are compared with those from the standard Round Robin scheduling algorithm, while the other simulation parameters remain the same (for more information about the simulation parameters refer to “Scenario” in section III). Furthermore, the performance of the new scheduler is evaluated for different bandwidth sizes and traffic share percentages. Each scenario has been run ten consecutive times in order to have a sufficient number of random users’ realizations.

1. Random traffic share and 10MHz bandwidth

The plots below provide a comparison between the performance of the Priority Scheduler and Standard Round Robin scheduler in the case of a random traffic share and a bandwidth size of 10MHz.

A) 100 users and 300 time slots



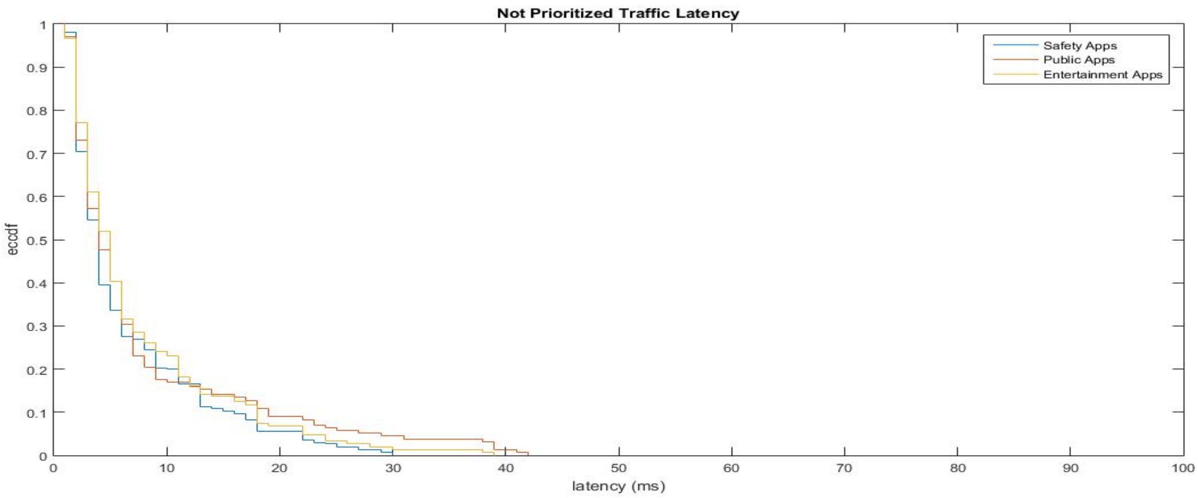


Fig. 22 Traffic latency with the Priority Scheduling algorithm (top) and Standard Round Robin scheduler for 100 users and 300 time slots

In Fig. 22 we can see a comparison between the two schedulers in the case of 100 users and 300 timeslots. In the results from the Priority Scheduler (PS), the traffic is prioritized as expected, while in the Standard Round Robin (SRR) the different types of traffic have almost the same latency values.

In Table 6 a summarize of the results for the two schedulers is shown.

	Priority Scheduler (PS)			Standart Round Robin Scheduler (SRR)		
	Safety Apps	Public Apps	Entertainment Apps	Safety Apps	Public Apps	Entertainment Apps
	latency(ms)			latency(ms)		
mean	4,98	6,63	10,1	6,34	7,5	7,25
max	35	60	67	30	42	39
10% of users	2	2	3	2	2	2
50% of users	3	4	5	4	4	5
90% of users	15	16	20	15	19	18
95% of users	15	26,3	34	22	28,7	22

Table 6. Summarized results for 100 users

We can see that for the Priority Scheduler results, the average and maximum latency for Safety Apps' traffic are 2 times lower than the Entertainment Apps' traffic. For PS results the average value of the latency for Safety and Public Apps is lower compared to the SRR results. Also, we can see that 95% of the Safety Apps' traffic latency is 15ms or less, while for SRR this value is 22ms.

Fig 23. shows a comparison between the different types of traffic for each scheduler. Due to the prioritization in PS, Safety and Public Apps' latency is always lower compared the SRR, while the Entertainment Apps' latency is always higher.

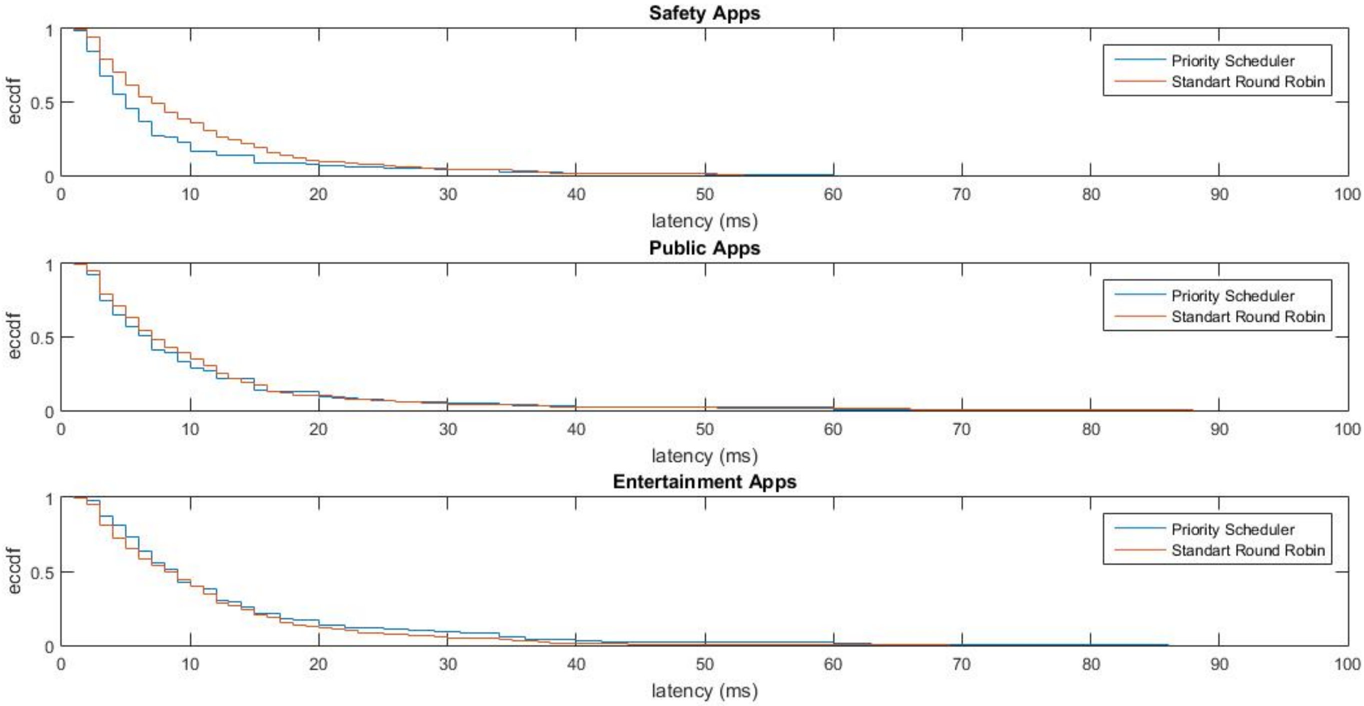
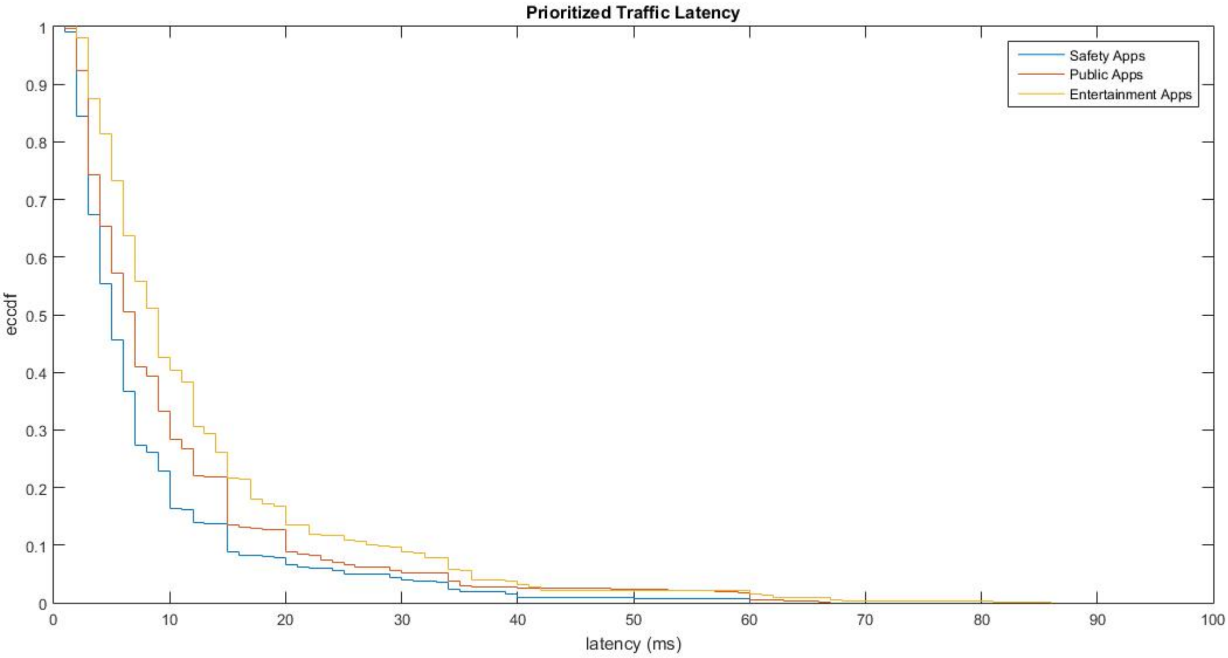


Fig. 23. Applications' traffic latency for 100 users

B) 200 users and 300 time slots

For the simulation with 200 users and 300 time slots, a similar relation can be observed. For SRR scheduler the values of the different types of traffic are even closer and in the range of 30 to 40ms they are almost the same (Fig 24).



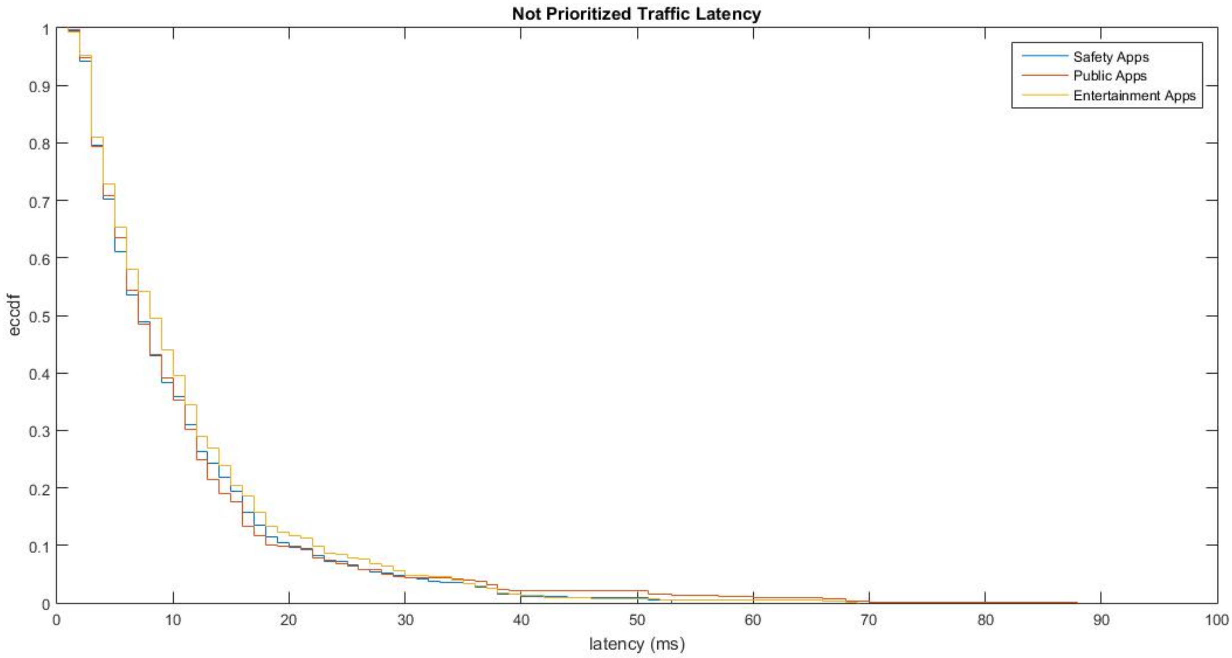


Fig. 24 Traffic latency with the Priority Scheduling algorithm (top) and Standard Round Robin scheduler for 200 users and 300 time slots

Table 7 shows the summarized results for both schedulers. Again for PS results the average value of the latency for Safety and Public Apps is lower compared to the SRR results. However we can see that for the Public Apps' latency the difference is really low, which can also be observed on Fig.25.

	PS			SSR		
	Safety Apps	Public Apps	Entertainment Apps	Safety Apps	Public Apps	Entertainment Apps
	latency(ms)			latency(ms)		
mean	7,78	9,97	12,45	10,19	10,38	10,9

max	60	67	86	53	88	69
10% of users	2	3	3	3	3	3
50% of users	5	7	9	7	7	8
90% of users	15	20	28	20	18,5	22
95% of users	27,56	34	36	29	28,5	30

Table 7. Summarized results for 200 users

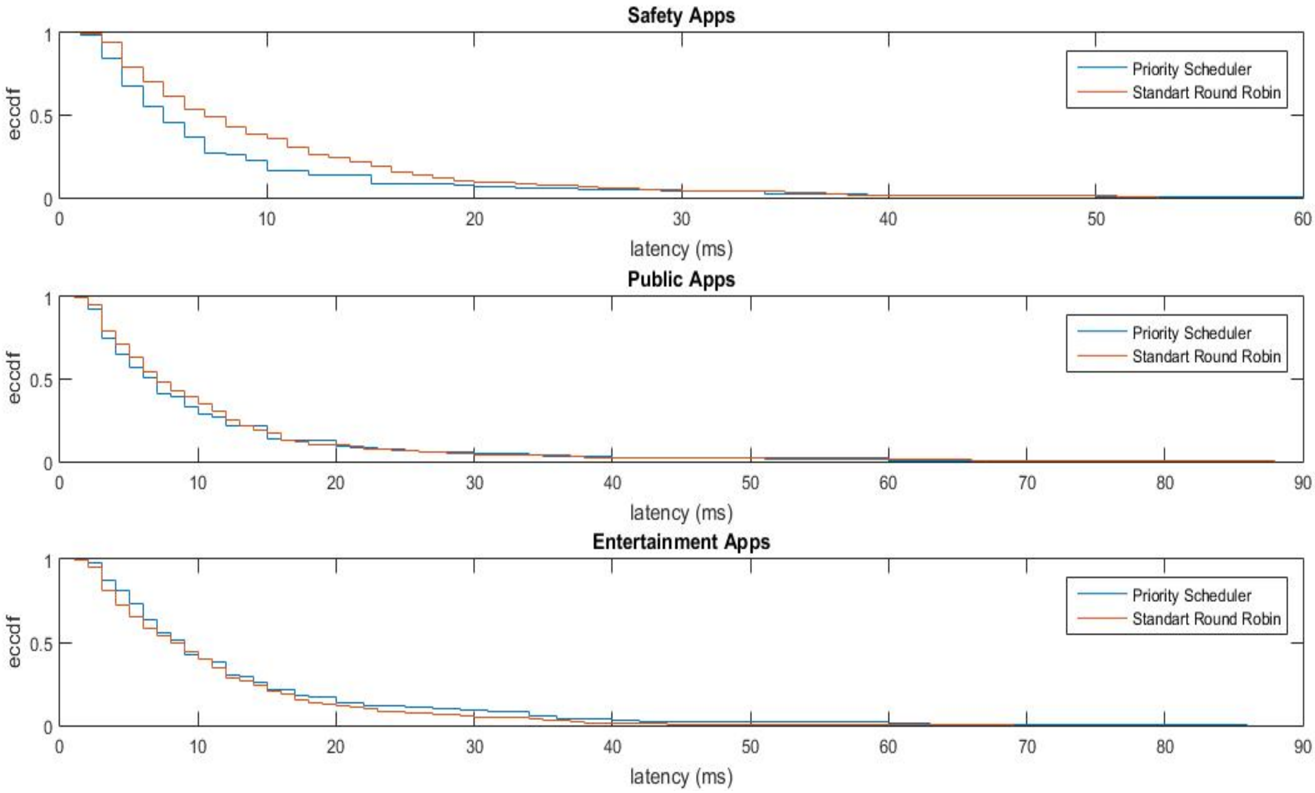


Fig. 25. Applications' traffic latency for 200 users

C) 500 users and 300 time slots

Fig 24 shows the results for 500 users and 300 time slots. Priority Scheduler plot is similar to the plots discussed in the previous simulations. However, the prioritization for the Public Apps' traffic is not so efficient. In Fig 25, we can see that in the range 48-70ms, Entertainment Apps' latency is lower than the Public Apps' latency.

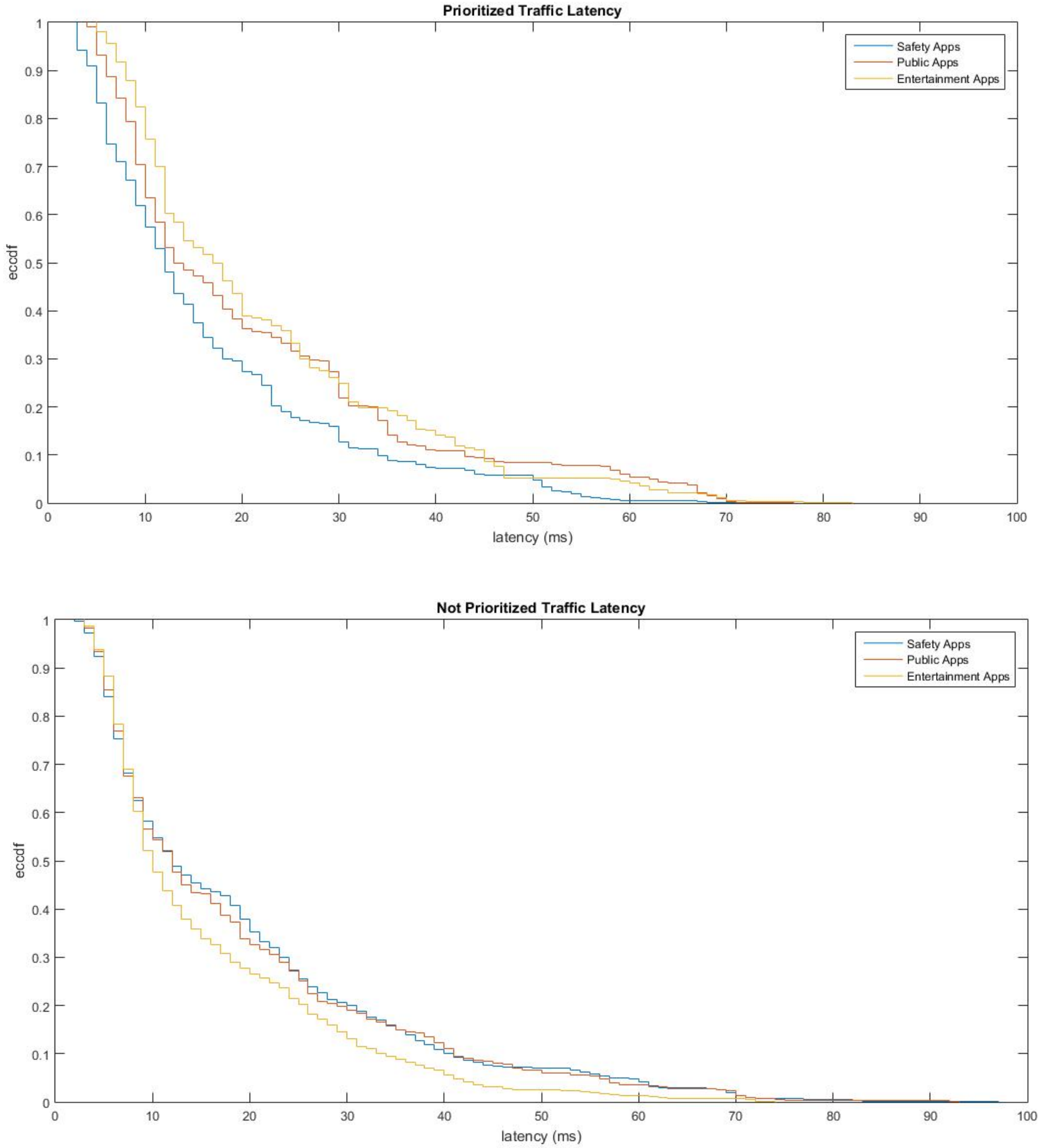


Fig. 26 Traffic latency with the Priority Scheduling algorithm (top) and Standard Round Robin scheduler for 500 users and 300 time slots

The prioritization of the Safety Apps' traffic had resulted in higher latency values for the Public Apps' traffic, which can be seen in Table 8 and Fig 27. For the PS results, the average latency for Public Apps is higher than the one for the SRR. Also, we can see similar differences between the percentiles. For example, in SRR results, 95% of the users had achieved latency of 56ms, while for PS results this value is higher (62ms).

	PS			SSR		
	Safety Apps	Public Apps	Entertainment Apps	Safety Apps	Public Apps	Entertainment Apps
	latency(ms)			latency(ms)		
mean	16,2	20,73	21,75	19	18,6	15,7
max	77	77	83	97	93	74
10% of users	5	6	8	5	5	5
50% of users	12	13	16	12	12	10
90% of users	34	43	45	41	41	33,3
95% of users	50	62	58	58	56	41

Table 8. Summarized results for 200 users

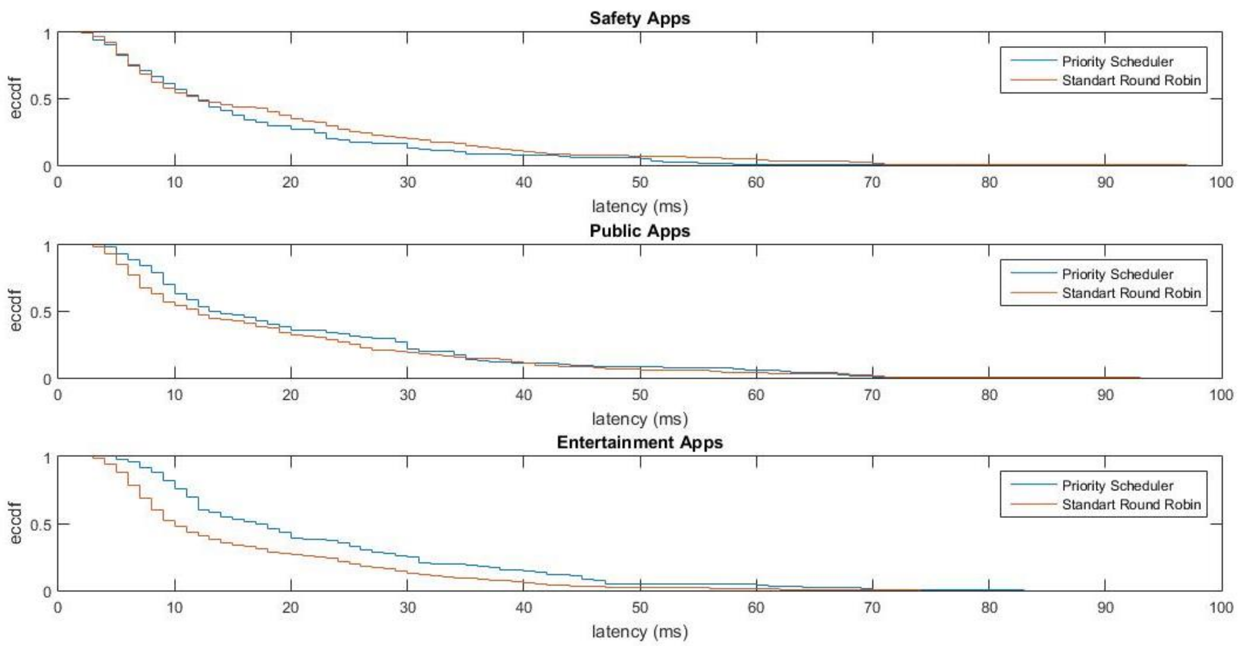


Fig. 27. Applications' traffic latency for 500 users

Since the user density in the network is higher for the following simulation, a significant amount of discarded packets is observed. Table 9, provides information about the discarded packets for the last 5 simulations.

	PS					SRR				
Simulation No.	Safety Apps	Public Apps	Entertainment Apps	Transmitted messages	Discarded messages	Safety Apps	Public Apps	Entertainment Apps	Transmitted messages	Discarded messages
	Number of discarded messages					Number of discarded messages				
1	6	7	14	831	27	9	9	5	830	23
2	3	7	13	863	23	8	10	17	835	35
3	6	10	12	829	28	5	5	3	884	13
4	6	10	12	842	28	10	7	7	835	24
5	6	7	14	831	27	5	5	3	884	13
Total	27	41	65	4196	133	37	36	35	4268	108

Table 9. Number of discarded packets for the last 5 simulations.

Due the prioritization in PS, the total number of discarded packets is higher compared to SRR. However, the number of discarded Safety Apps' packets is lower, which can be observed on Fig 28 too.

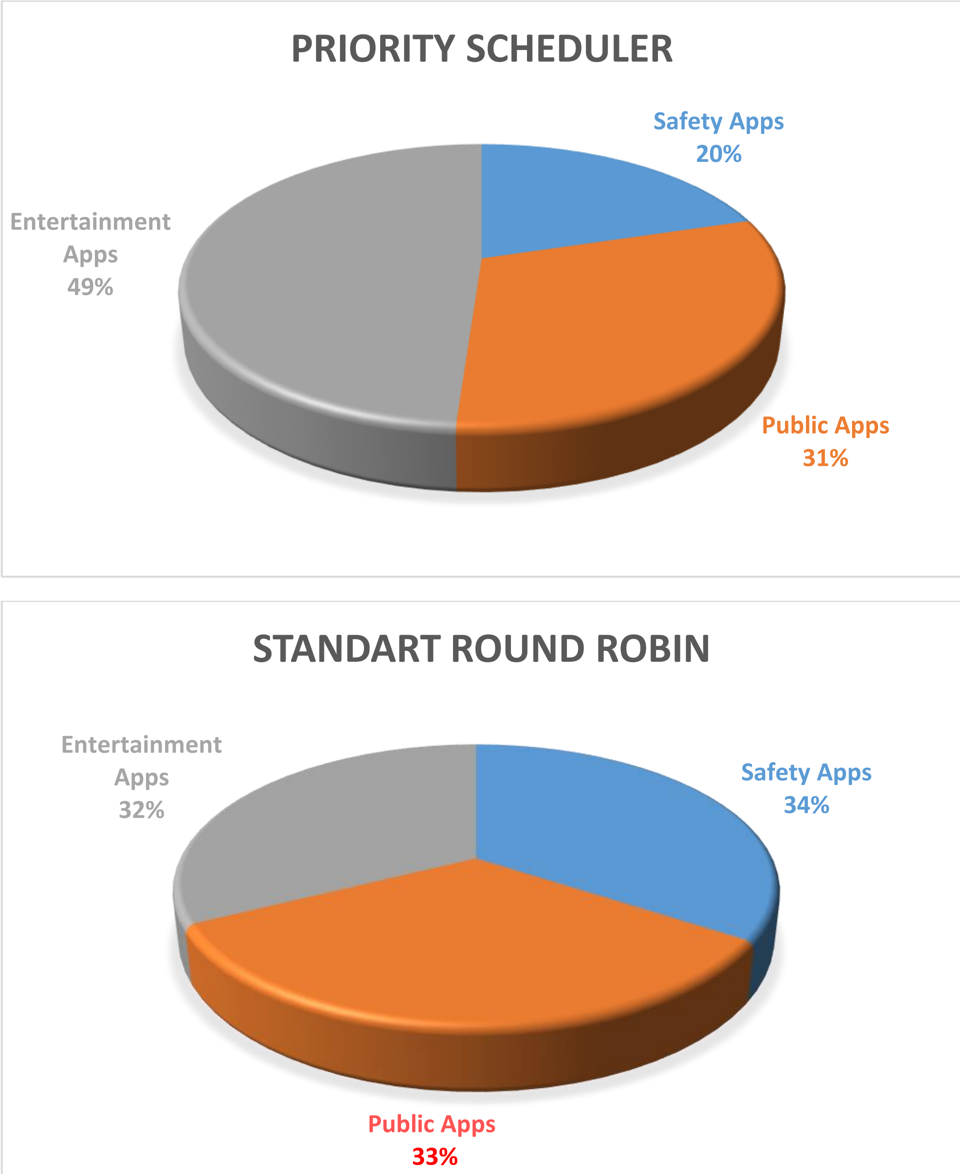


Fig. 28. Percentage of discarded packets per application

For PS around half of the discarded packets are from Entertainment traffic, while in SRR each application has almost equal share of discarded packets. Also we can see that for PS the percentage of discarded safety traffic is 14% lower compared to SRR. There is a slight improvement for public traffic too.

2. Random traffic share and different bandwidth size.

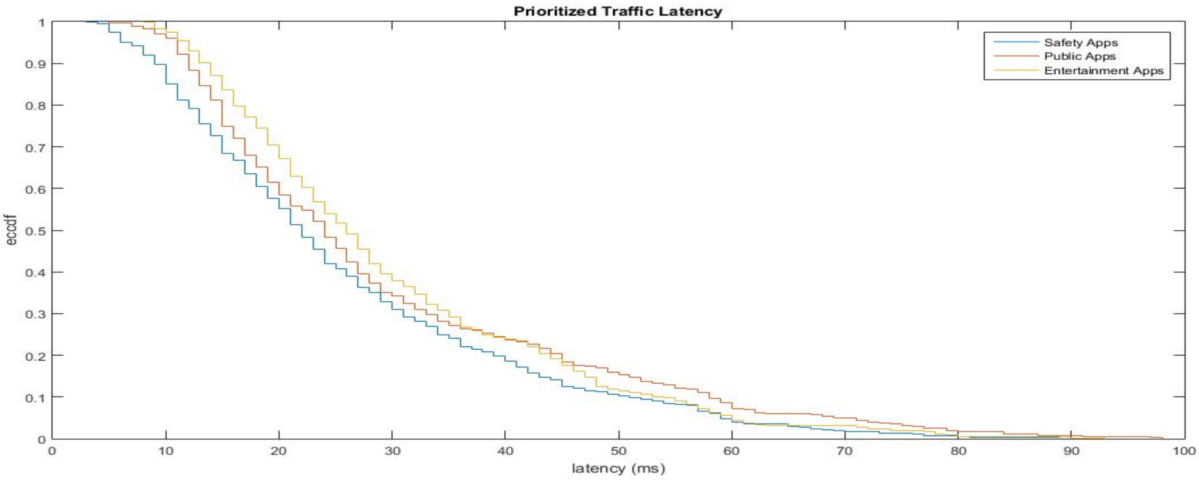
In this part, we will evaluate the performance of the priority scheduler in the case of different bandwidth sizes. The simulations have been run with 500 users, 300 time slots and bandwidth sizes of 5MHz, 10MHz, and 15MHz. The other goal is to evaluate the latency as a function of the system bandwidth.

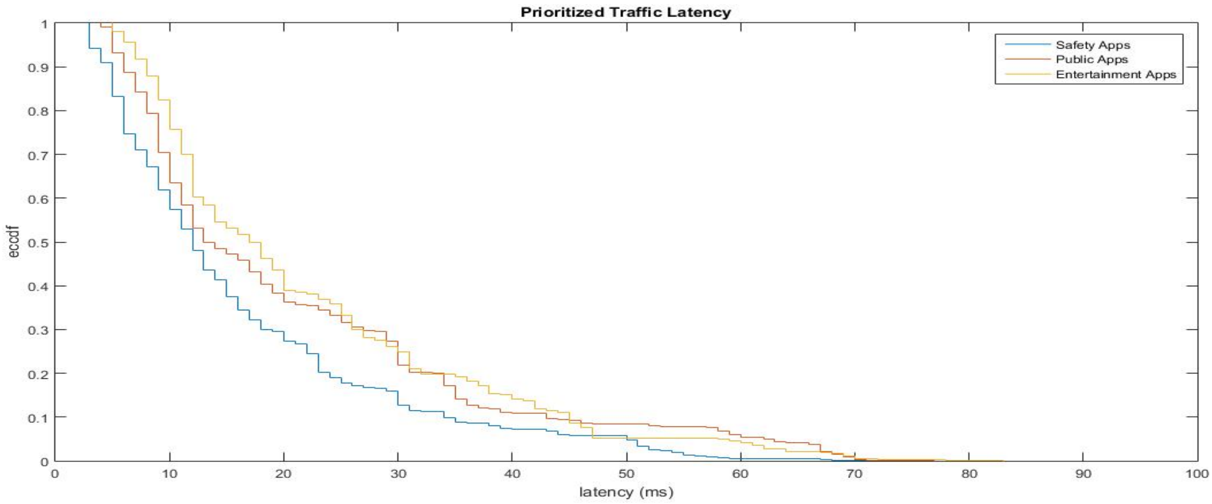
Table 9 shows the summarized results for each bandwidth

	5MHz			10MHz			15MHz		
	Safety Apps	Public Apps	Entertainment Apps	Safety Apps	Public Apps	Entertainment Apps	Safety Apps	Public Apps	Entertainment Apps
	latency (ms)								
min	3	5	8	2	2	2	2	2	3
mean	26,17	29,64	30,1	16,2	20,73	21,75	11,23	14,46	17,27
max	90	98	93	77	77	83	71	70	72
10%	9	12	14	5	6	8	3	4	6
50%	22	24	26	12	13	16	8	12	13
90%	51	58	53,9	34	43	45	21	30	39
95%	59	69,7	60	50	62	58	34	39	40

Table 9. Summarized results for 500 users and different bandwidth size.

The difference between the mean latency for each bandwidth size is in the range of 6-10ms. We can see that the prioritization works in all of the cases, but there are still some deviations.





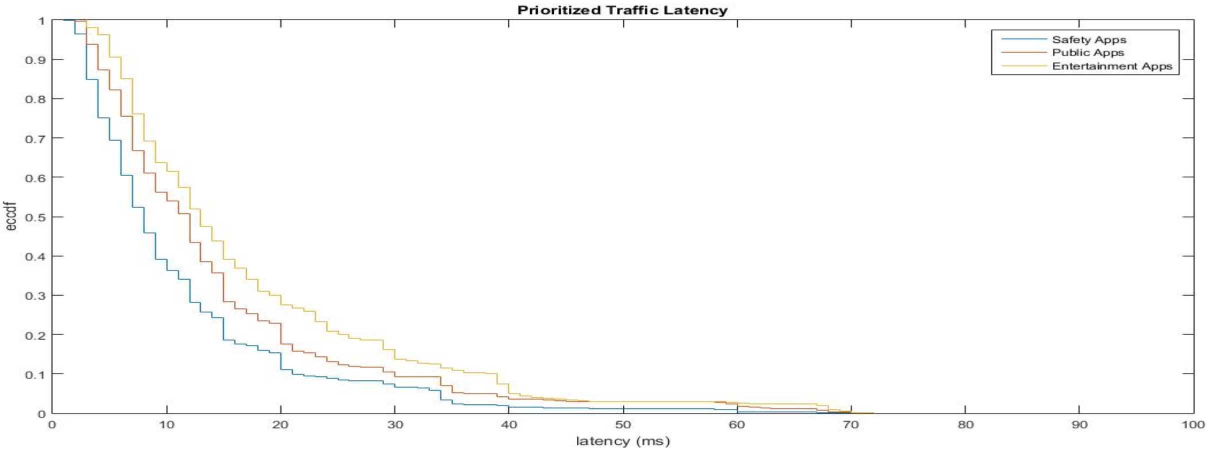


Fig.29. *Traffic latency for different bandwidth. 5MHz(top), 10MHz(middle) and 15MHz(bottom)*

For 5 and 10MHz we can see that there are deviations between the Public and Entertainment Apps' latency. The deviations occur in the same range and they are probably due to discarded Entertainment Apps' messages. For 15MHz these deviations also exist but they are not so drastic.

V. Conclusions

According to the received results, both PS and SRR are able to transfer the packets within the latency limit of 100ms. However, PS has better performance when it comes to latency sensitive traffic. In all simulation cases, the average Safety Apps' latency for PS is lower compared to the one in SRR. For PS the amount of discarded Safety Apps' packets is also lower, which is extremely important for vehicular communications, since those apps are providing road safety information.

In low user density networks, there is a slight improvement for Public Apps too. But in the case of 500 users the performance of PS when it comes to Public Apps' traffic is worse compared to SRR. Also in all cases the prioritization of Safety and Public Apps, results to worsening in the performance of the Entertainment Apps.

References:

1. Arena, F.; Pau, G. An Overview of Vehicular Communications. *Future Internet* **2019**, *11*, 27.
2. Emmanuel Ndashimye , Sayan K. Ray , Nurul I Sarkar Jairo A. Guti«errez , Vehicle-to-Infrastructure Communication over Multi-Tier Heterogeneous Networks: A Survey, Computer Networks (2016)
3. Qiong Yang, Lin Wang, Weiwei Xia, Yi Wu, Lianfeng Shen, Development of On-Board Unit in Vehicular Ad-Hoc Network for Highways (2014)
4. https://www.its.dot.gov/pilots/cv_pilot_apps.htm
5. http://www.mogi.bme.hu/TAMOP/jarmurendszerek_iranyitasa_angol/math-ch09.html
6. WeiXiang· KanZheng, Xuemin(Sherman)Shen Editors. 5G Mobile Communications (2017)
7. K. Zheng, Q. Zheng, P. Chatzimisios, W. Xiang, Y. Zhou, Heterogeneous Vehicular Networking: A Survey on Architecture, Challenges, and Solutions. IEEE Commun. Surv. Tutorials, 17(4), 2377–2396, (2015) Fourthquarter
8. W. Saad et al., Coalitional game theory for communication networks. IEEE Signal Process. Mag. 26(5), 77–97 (2009)
9. Y. Li, D. Jin, J. Yuan, Z. Han, Coalitional games for resource allocation in the device-to-device uplink underlaying cellular networks. IEEE Trans. Wirel. Commun. 13, 3965–3977 (2014)
10. Z. Zhou, M. Dong, K. Ota, R. Shi, Z. Liu, T. Sato, Game-theoretic approach to energy-efficient resource allocation in device-to-device underlay communications. IET Commun. 9(3), 375–385, (2015)
11. Kais Elmurtadi Suleiman, Otman Basir, "Adaptive incentive-compatible routing in VANETs for highway applications", *Information Technology Electronics and Mobile Communication Conference (IEMCON) 2017 8th IEEE Annual*, pp. 670-679, 2017.
12. Martin Klaus Müller, Fjolla Ademaj, Thomas Dittrich, Agnes Fastenbauer, Blanca Ramos Elbal, Armand Nabavi, Lukas Nagel, Stefan Schwarz, Markus Rupp, "Flexible multi-node simulation of cellular mobile communications: the Vienna 5G System Level Simulator"
13. 3rd Generation Partnership Project; Technical Specification Group Radio Access Network; Study on LTE-based V2X Services; (Release 14)