UNIVERSITAT POLITÈCNICA DE CATALUNYA BARCELONATECH

Doctoral Thesis

Contribution to the Traffic Engineering in Wireless Mesh Networks.

by:

Juan Pablo Astudillo León

Ph.D. Advisor:

Dr. Luis J. de la Cruz Llopis

Thesis submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Telematics

SISCOM (Smart Services for Information Systems and Communication Networks)

Departament of Network Engineering

April 29, 2020

Agradecimientos

El desarrollo de este proyecto no habría sido posible sin el apoyo y el estímulo de Luis, bajo cuya supervisión fue posible el desarrollo del proyecto, mas que mi director lo que considero un gran amigo. Agradezco mucho las charlas, consejos recibidos, momentos en familia, y también su contención en los momentos difíciles. Espero y seguro que seguiremos colaborando juntos, por que como siempre dice, hay trabajo por hacer.

Además quiero agradecer a las siguientes personas que formaron parte de mi familia acá en Barcelona, que como siempre lo he dicho lo he vivido en etapas. La primera etapa en mis primeros días acá en Barcelona agradezco a mis grandes amigas Ali y Meli, con las cuáles vivimos gratos y valiosos momentos en cada rincón de esta hermosa ciudad. Sobretodo con Alicia, que comenzamos igual nuestros estudios doctorales, y muchas cosas las vivimos iguales. En la segunda etapa, donde ya comienza mi vida a tonarse mas familiar, en la cuál Nathaly, Christian y sobretodo Nicole se sumaron. Con Christian y Nicole hicimos un gran equipo. Las gratas charlas con Christian que es una persona con la cuál disfruto mucho conversar y de esas experiencias aprender un poco más. Te considero uno de mis mejores amigos. Nicole fue y es aún una parte fundamental en mi persona, de la cuál agradezco las innumerables experiencias vividas y disfrutadas. Los viajes y ocurrencias de Nicci lo llevo muy profundo en mi corazon. La tercera parte de mi vida doctoral, se sumaron personas muy especiales como Víctor, Gustavo, Byron, Layla y David, mis agradecimientos también para ustedes. Las parrilladas del golden, nuestros inventos en la cocina y gratos momentos compartidos en Riera Blanca con David y especial con Víctor los almaceno con mucho cariño. Además, las escapadas y grandes momentos compartidos con Layla, y las ocurrencias de mi amigo Byron, que es una persona que puede conversar de todo y con una gran seguridad, mis agradecimientos y cariño también para ustedes. En mi última etapa agradezco Anita, Luis, Pablo y José, con los cuales compartimos gratos momentos. Como no olvidar ademas las charlas en Granollers con el Pupi y Joseph, los estimo mucho mis amigos. No obstante, quiero agradecer de manera muy especial a Leti por todas las experiencias vividas, de las cuales he aprendido mucho de ella, y en especial de su hermosa cultura mexicana. Además, las grandes personas que conocí gracias a ella. Todo este camino recorrido y los gratos momentos al final de nuestra etapa doctoral para mi es muy especial, y no lo puedo describir en palabras pero de seguro no lo olvidaré.

I would like to express my sincere gratitude to Anthony Busson and Thomas Begin for having made possible my stay in the LIP laboratory of the Ecole Normale Supérieure (ENS) in Lyon, France. This experience was very enriching, since part of Chapter 4 includes much of the work we did in Lyon and the collaborations carried out from Barcelona. I am very happy to have part of my work published with you, and I hope to continue collaborating with you. The experience lived with you in Lyon was very enriching from a personal and investigative point of view.

Mi agradecimientos también a mis amigos del laboratorio: Nitin, Akram, Godfrey, Khaled, Ahmed, Ahmad. Mis agradecimientos también al grupo de SISCOM, en especial a Mónica por todo el apoyo y ayuda brindado. Además,

Asimismo quiero reconocer el apoyo brindado por la Secretaría de Educación Superior, Ciencia, Tecnología e Innovación (Senescyt), institución que financió mediante una beca mis años de estudios e investigación en la Universitat Politècnica de Catalunya. Ademas quiero agradecer a los proyectos INRISCO (TEC2014-54335-C4-1-R) y MAGOS (TEC2017-84197-C4-3-R) que financiaron también este trabajo.

También quiero agradecer a los miembros del jurado que aceptaron revisar el resultado del presente trabajo.

No puedo terminar sin agradecer a mi familia, en cuyo estímulo constante y amor he confiado a lo largo de mis años estudiantiles...

Abstract

Nowadays, we live in a modern society in which people and devices are interconnected anywhere and anytime. Under this premise, both the infrastructure and the services offered have evolved and diversified in a drastic way. In fact, many of these services are transported in different networks types. Decentralized networks (or without infrastructure) are being widely used to support these services. They allow greater accessibility for users due to a large number of advantages. For instance, self-creation, self-configuration, easy installation in areas of difficult access, maintenance and scalability make that kind of networks attractive to service providers. Among them, Wireless mesh networks are decentralized networks that have been widely studied in different research areas such as community networks, disaster scenarios, public safety and surveillance. Besides, these network types are more structured than the traditional wireless ad hoc networks, and thus, they can support more complex protocols.

Wireless mesh networks have been also studied and evaluated in the Smart Grid scenario. Smart Grids are a new paradigm in which the traditional electricity transport infrastructures are addressed. In this context, the electricity network is not longer focused only on the generation, distribution and transport of electricity to subscribers. Now, it is a robust network that includes a data communication network. The objective of having a data communication network together with the electrical is to provide an efficient service from the control center to the user as well as to give feedback of the correct operation of the electricity and data networks to the control center. As the electrical transport infrastructure, the associated data network is divided in different subnetworks. This thesis is mainly focused on the improvement of the performance of one of those subnetworks, the so-called Smart Grid Neighborhood Area Network. The contributions focus on improving the data routing, providing traffic differentiation with Quality of Service (QoS) provision, congestion control mechanisms, an emergency system which deals anomalous network situations and fair distribution of network resources.

Several applications are transmitted upstream from the users to the control center, as well as downstream from the control center towards the users. In general, upstream communication involves tasks such as meter reading, billing data or electricity consumption, while downstream communication allows the smart grid to take actions in different network situations such as power peaks or emergency situations. In the first part of the thesis, the work is focused on improving the routing mechanism. To do this, a multipath routing mechanism is proposed, where the traffics that are most important are transmitted over the best communication links, while the lowest priority traffics are transmitted over the paths with less reputation (less routing metric). In order to improve even more the benefits obtained, a multichannel scheme is proposed to separate both control traffic and data traffic, and use the less congested channels to transmit the most priority traffic types.

Smart Grids offer many services and some of them are very demanding in terms of QoS. Therefore, infrastructure failures, attacks and high congestion situations can greatly reduce the network performance. In order to face with these problems, the network must be able to offer a minimum quality of service to the most priority applications handling some traffic control techniques. With this goal in mind, in this thesis some congestion control mechanisms are also proposed. In the first of these mechanisms, the decision of whether a packet should be retransmitted or not is made in a distributed and independent way by each one of the network nodes, depending on the network conditions (mainly the averaged channel utilization and the buffers occupancy) that the node itself is observing. That is, an intermediate node can directly drop a data packet if it observes that the transmission channel is being used above a certain threshold. This mechanism considers again the existence of traffics with different priorities, so

that, less priority traffic has a higher probability of being discarded. Furthermore, an emergency system is coupled to the congestion control mechanism. With this strategy, the NAN is able to take global actions (in a short time) to face anomalous situations providing with even more transmission probability for traffics with higher QoS requirements. To this end, a emergency signalling that can be triggered automatically or manually is also proposed.

A fair distribution of network resources is also an important research field in the Smart Grid. Keep in mind that in this scenario, the nodes are static and each of them transmits upstream data flows to the data concentrator. Therefore, depending on their geographical location, some nodes may be more favored than others. Besides, some nodes can monopolize the network resources if they are not regulated. For this reason, in this thesis another distributed congestion control algorithm is proposed that runs in each node. The objective here is to provide a fair distribution of network resources regardless of the geographical position and the transmission rate. That is, all the nodes will have the same opportunities to transmit their data to the control center. The proposed solution is agnostic to the network, mac and physical layers.

The last contribution made with this thesis is focused on the application of machine learning techniques to obtain again a better performance of the data networks under study. In this sense, a new congestion control mechanism is proposed, which, like the previous ones, provides different quality of service to data flows with different priorities. For this, a complete framework is proposed, including the generation, preprocessing and evaluation of the data necessary for the training of the machine learning algorithms that will be used. The proposal is also implemented and evaluated in the Smart Grid NANs environment.

Resumen

Hoy en día, vivimos en una sociedad moderna en la que las personas y los dispositivos están interconectados en cualquier lugar y en cualquier momento. Bajo esta premisa, tanto la infraestructura como los servicios ofrecidos han evolucionado y diversificado de manera drástica. De hecho, muchos de estos servicios se transportan en diferentes tipos de redes. Las redes descentralizadas (o sin infraestructura) se están utilizando ampliamente para soportar estos servicios. Permiten una mayor accesibilidad para los usuarios debido a una gran cantidad de ventajas. Por ejemplo, la creación automática, la configuración automática, la instalación fácil en áreas de difícil acceso, mantenimiento y escalabilidad hacen que este tipo de redes sea atractivas para los proveedores de servicios. Entre ellas, las redes de malla inalámbricas son redes descentralizadas que han sido ampliamente estudiadas en diferentes áreas de investigación, como redes comunitarias, escenarios de desastres, seguridad pública y vigilancia. Además, estos tipos de red son más estructurados que las redes ad hoc inalámbricas tradicionales y, por lo tanto, pueden admitir protocolos más complejos.

Las redes de malla inalámbricas también se han estudiado y evaluado en el escenario de redes eléctricas inteligentes. Las redes eléctricas inteligentes son un nuevo paradigma en el que se abordan las infraestructuras tradicionales de transporte de electricidad. En este contexto, la red eléctrica ya no se centra solo en la generación, distribución y transporte de electricidad a los suscriptores. Ahora, es una red robusta que incluye una red de comunicación de datos. El objetivo de tener una red de comunicación de datos junto con la eléctrica es proporcionar un servicio eficiente desde el centro de control al usuario, así como dar retroalimentación sobre el correcto funcionamiento de las redes de electricidad y datos al centro de control. Como la infraestructura de transporte eléctrico, la red de datos asociada se divide en diferentes subredes. Esta tesis se centra principalmente en la mejora del rendimiento de una de esas subredes, la llamada red de área de vecindad de las redes eléctricas inteligentes. Las contribuciones se centran en mejorar el enrutamiento de datos, proporcionando una diferenciación del tráfico con la provisión de calidad de servicio (QoS), mecanismos de control de congestión, un sistema de emergencia que trata situaciones anómalas de la red y una distribución justa de los recursos de la red.

Varias aplicaciones se transmiten desde los usuarios al centro de control, así como desde el centro de control hacia los usuarios. En general, la comunicación hacia el centro de control implica tareas como la lectura de medidores, los datos de facturación o el consumo de electricidad, mientras que la comunicación hacia los suscriptores permite que la red eléctrica inteligente tome medidas en diferentes situaciones de la red, como picos de energía o situaciones de emergencia. En la primera parte de la tesis, el trabajo se centra en mejorar el mecanismo de enrutamiento. Para hacer esto, se propone un mecanismo de enrutamiento de múltiples rutas, donde los tráficos que son más importantes se transmiten a través de los mejores enlaces de comunicación, mientras que los tráficos de menor prioridad se transmiten a través de las rutas con menos reputación (menos métrica de enrutamiento). Para mejorar aún más los beneficios obtenidos, se propone un esquema multicanal para separar tanto el tráfico de control como el tráfico de datos, y utilizar los canales menos congestionados para transmitir los tipos de tráfico más prioritarios.

Las redes eléctricas inteligentes ofrecen muchos servicios y algunos de ellos son muy exigentes en términos de QoS. Por lo tanto, las fallas de infraestructura, los ataques y las situaciones de alta congestión pueden reducir en gran medida el rendimiento de la red. Para enfrentar estos problemas, la red debe poder ofrecer una calidad de servicio mínima a las aplicaciones más prioritarias mediante algunas técnicas de control de tráfico. Con este objetivo en mente, en esta tesis también se proponen algunos mecanismos de control de congestión. En el primero de estos

mecanismos, cada uno de los nodos de la red decide de manera distribuida e independiente si un paquete debe o no ser retransmitido, dependiendo de las condiciones de la red (principalmente la utilización promedio del canal y la ocupación de los buffers) que el nodo mismo está observando. Es decir, un nodo intermedio puede descartar directamente un paquete de datos si observa que el canal de transmisión se está utilizando por encima de un cierto umbral. Este mecanismo considera nuevamente la existencia de tráficos con diferentes prioridades, de modo que, el tráfico menos prioritario tiene una mayor probabilidad de ser descartado. Además, un sistema de emergencia está acoplado al mecanismo de control de congestión. Con esta estrategia, la NAN puede tomar acciones globales (en poco tiempo) para enfrentar situaciones anómalas, lo que proporciona aún más probabilidad de transmisión para tráficos con mayores requisitos de QoS. Con este fin, también se propone una señalización de emergencia que puede activarse automática o manualmente.

Una distribución justa de los recursos de la red también es un campo de investigación importante en las redes eléctricas inteligentes. Tenga en cuenta que en este escenario, los nodos son estáticos y cada uno de ellos transmite flujos de datos hacia al concentrador de datos. Por lo tanto, dependiendo de su ubicación geográfica, algunos nodos pueden ser más favorecidos que otros. Además, algunos nodos pueden monopolizar los recursos de la red si no están regulados. Por esta razón, en esta tesis se propone otro algoritmo de control de congestión distribuido que se ejecuta en cada nodo. El objetivo aquí es proporcionar una distribución justa de los recursos de la red, independientemente de la posición geográfica y la velocidad de transmisión. Es decir, todos los nodos tendrán las mismas oportunidades para transmitir sus datos al centro de control. La solución propuesta es independiente de la red, mac y capas físicas.

La última contribución realizada con esta tesis se centra en la aplicación de técnicas de aprendizaje automático para obtener nuevamente un mejor rendimiento de las redes de datos en estudio. En este sentido, se propone un nuevo mecanismo de control de congestión que, al igual que los anteriores, proporciona diferente calidad de servicio a los flujos de datos con diferentes prioridades. Para esto, se propone un marco completo, que incluye la generación, el preprocesamiento y la evaluación de los datos necesarios para la capacitación de los algoritmos de aprendizaje automático que se utilizarán. La propuesta también se implementa y evalúa en el entorno de Smart Grid NANs.

Acronyms

AHP	Analytical Hierarchy Process
AIFS	Arbitration Interframe Space Number
AIFSN	Arbitration Interframe Space
ALM	Airtime Link Metric
AMI	Advanced Metering Infrastructure
AODV	Ad hoc On-Demand Distance Vector Routing
CBR	Constant Bit Rate
CN	Community Network
CW	Contention Window
DCF	Distributed Coordination Function
DIFS	DCF Interframe Space
DT	Decision Tree
EA-HWMP	Emergency Aware Congestion Control - HWMP
EDCA	Enhanced Distributed Channel Access
EPT	Expected Path Throughput
EV	Electric Vehicle
ETX	ExpectedTransmission Count
EWMA	Exponentially-Weighted Moving Average
FDCC	Fair and Distributed Congestion Control
HAN	Home Area Network
HWMP	Hybrid Wireless Mesh Protocol
IAETT	Interference Aware Expected Transmission Time
MAC	Medium Access Control
MBSS	Mesh Basic Service Set
MCCA	MCF Controlled Channel Access
MD	Minimum Delay
MCF	Mesh Coordination Function
ML	Minimum Loss
MLCC	Machine Learning Congestion Control
MPC-HWMP	Multi-Path Multi-Channel Hybrid Wireless Mesh Protocol
MSTA	Mesh Station
MPM	Mesh Peering Management Protocol
MPR	MultipointRelay nodes
NAN	Neighborhood Area Network
NHDP	Neighborhood Discovery Protocol
OGM	Originator Message
OLSR	Optimized Link State Protocol
PLC	Power Line Communication
pdf	probability distribution function

PDR	Packet Delivery Ratio
PREP	Path Reply
PREQ	Path Request
PERR	Path Error
QoS	Quality of Service
ROC	Receiver Operating Characteristic
RREQ	Route Request
RREP	Route Reply
RLMTs	Relevant Link Metric Types
SM	Smart Meter
SN	Sequence Number
STA	Station
TXOP	Transmission Opportunity
UDP	User Datagram Protocol
WAN	Wide Area Network
WANET	Wireless Ad hoc Network
WLAN	Wireless Local Area Network
WMN	Wireless Mesh Netowork

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Today's society involves people and devices of all kinds connected all the time [4], which depends on services offered online and that are transmitted by different technologies. In such scenario, Wireless Mesh Networks (WMNs) have evolved as a cost effective possible solution for the uninterrupted access of users to networking facilities. Valued features like robustness, reliability, resilience, easy deployment and maintenance, self-forming and self-configuration, make WMNs an important alternative to achieve an always-on connectivity.

Wireless Mesh Networks (WMNs) are based on the ad hoc networking paradigm [5], and they provide self-configurable and self-healing characteristics to dynamic network topologies [6]. Nevertheless, WMNs are more structured than ad hoc networks and can support more complex protocols and routing metrics [7–14]. Furthermore, WMNs allow the interconnection with other types of networks such as wireless local area networks, cellular networks, metropolitan area networks, vehicular networks [15], sensor networks, disaster scenarios [16–19], smart grid networks, and personal and body area networks [20, 21] and the global Internet [22, 23].

Figure 1.1 presents the WMNs architecture where three network types can be seen: infrastructure meshing, client meshing, and hybrid WMNs.

- Infrastructure meshing provides a wireless backbone to the clients. This backbone is made up of static mesh routers.
- Client meshing provides end user applications and routing functionalities to client nodes.
- A combination of infrastructure and client meshing can be used with the hybrid WMN [24].

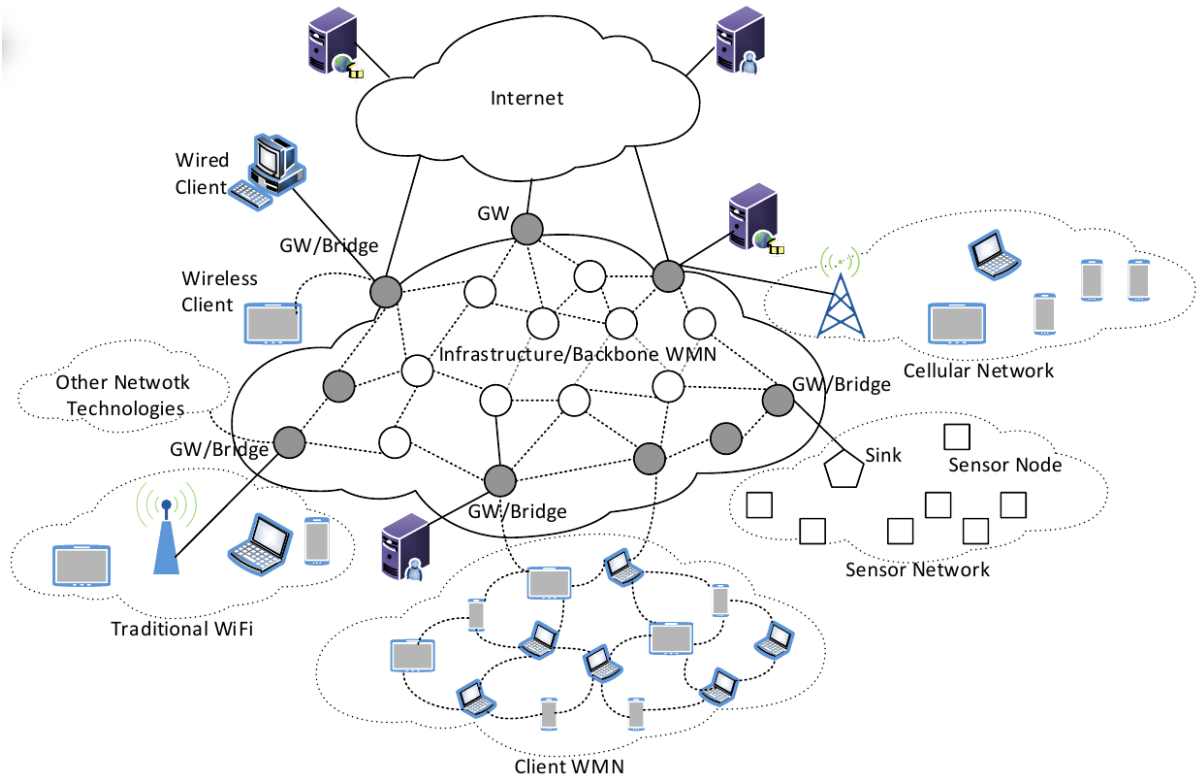


FIGURE 1.1: Wireless Mesh Network paradigm [24].

WMNs exhibit self-healing mechanisms, since a failure node does not imply the network failure. Self-configurable characteristics allows to calculate new optimal paths for routing. Furthermore, network coverage between two nodes is increased by using a multi-hop scheme. Besides, self-organization and self-configuration capacities allow a better growth and maintenance of these networks [25].

Given these characteristics, WMNs have been evaluated in many applications such as:

- **Community Networks:** Several routing protocols for WMNs are studied and evaluated in Community Networks (CNs). Community Networks as described in [26], are an open and distributed infrastructure where researchers can deploy experimental services, perform tests or access to open data. In [27], it is described the purposes and benefits for which the Community Networks were conceived.

  Representative community network examples are Freifunk (FF) [28] in Germany, the Athens Wireless Metropolitan Network (AWMN) [29] in Greece, FunkFeuer [30] in Austria, guifi [31] in Spain, and Ninux.org [32] in Italy.

- **Smart Grids (SGs):** The SGs are considered an intelligent network that comprise all the power systems stages such as generation, transmission and distribution. Besides, they consider different communication mechanisms between customers and service providers [33]. The objective is to optimize all the inherent functionalities from the generation of electricity to the final customer service. The communication structure between backhaul aggregation points to the core backbone utility center is carried over different types

of communication networks [34]. Communication network requirements for major smart grid applications are described in [35, 36].

- Other current applications for WMNs are: Internet of Things [37, 38], Blockchain [39, 40], Multimedia services [41–44], surveillance systems and public safety [45–47] and so on.

Several researchers have focused their efforts in order to improve the network performance of the Smart Grids, and specifically for Neighborhood Area Networks (NANs). The NANs are a part or Smart Grid data communication network. In this thesis, the work is focused on improving the network performance obtained by those NANs when the selected technology is a wireless multi-hop network,

In the following subsection an overview for SG is described.

1.1 Smart Grid data communication network

Electrical energy is currently an essential resource all around the world. With the increase in the use of new technologies in all sectors of human activity, it is easy to predict that the consumption of this type of energy will grow considerably in the near future. For this reason, great research and development efforts have been made, with the aim of improving the generation processes, transport networks and storage systems for this energy. The Smart Grid networks are the result of the work carried out to obtain improvements in the management, operation and maintenance of the transport infrastructure, as well as in the efficiency with which the energy is used. Its main objective could then be considered as achieving the best use of electrical energy through an improvement in the management and maintenance of the energy sources and the transport infrastructure. At the same time, new services are offered to both supplying companies and consumers.

With these objectives in mind, one of the main advances is being made in the improvement of the data networks associated with the electricity transport infrastructures. These data networks are responsible for carrying and delivering all the control, management, maintenance and security information of the electricity network infrastructure, as well as the applications that allow a better management of the available resources. The data networks are therefore a fundamental part of the Smart Grid, and so their reliability, availability and security have to be guaranteed in all situations, taking into account that they can stop providing correctly their service due to intrinsic (hardware, software, communications protocols, ...) or extrinsic (weather conditions, malicious agents, terrorist attacks, ...) failures [48].

Figure 1.2 presents the Smart Grid data communication network, which it is made up of several subnetworks. The different smart meters (SM), devices and other utilities present inside the customer homes are interconnected by the Home Area Network (HAN). These HANs are in turn interconnected through the Neighborhood Area Network (NAN), and finally the information can reach the control centers through a Wide Area Network (WAN). To implement all these subnetworks, different technologies can be chosen. Selectable technology standards for HANs can be, among others, IEEE 802.15.4 (Low-Rate Wireless Personal Area Networks, LR-WPAN) or IEEE 802.11 (Wireless Local Area Networks, WLAN). For NANs, Power Line Communication (PLC) technologies, or standards such as IEEE 802.15.4g or IEEE 802.11s (Wireless Mesh Networks), can be considered.

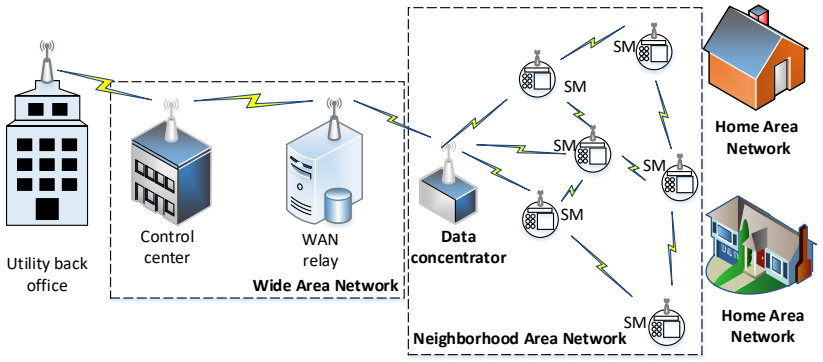


FIGURE 1.2: Smart Grid data subnetworks.

Guaranteeing that quality of service offered is the required by every application, while maintaining a high efficiency in the use of resources, is one of the most important issues to be taken into account both during the planning process and during the maintenance and operation of the network. It is therefore essential a deep knowledge of the services that will be provided. These issues are especially important when working with Smart Grid, given the critical importance of the power grid infrastructure. The offered services belongs to very different types, and therefore their service quality requirements are also different [49]. Generally speaking, most Smart Grid applications have strong security and reliability requirements. However, their bandwidth needs, as well as their behavior in high packet losses or high latency situations, can be very different. Thus, some applications, such as Substation Automation Systems (SAS) or Overhead Transmission Line Monitoring, present very strict requirements in terms of latency, but they are not as demanding in terms of bandwidth. On the other hand, Demand Response Management or Advanced Metering Infrastructure (AMI) generally consume more bandwith but can allow greater delays.

1.2 Objectives

The objectives of this dissertation are devoted to improve the network performance of wireless multi hop networks, and specifically the Smart Grid Neighborhood Area Networks. The main contributions of this thesis are described below:

- **Network routing and multichannel allocation.** Multi-path and multichannel allocation schemes together with traffic differentiation are proposed, implemented and evaluated for high network loads. For this, the default IEEE 802.11s routing protocol has been modified.
- **Congestion control mechanisms and emergency system.** Different solutions to deal with high network congestion situations together with traffic differentiation are also proposed, implemented and evaluated. For this purpose, three different mechanisms which are implemented individually at each node are presented. First, a congestion control mechanism which works together with an emergency system is considered. For this approach, the forwarding decisions are up to the relay nodes and these decisions are based on congestion control functions. Furthermore, control signaling to trigger emergency situations

is added to the proposal. Second, a distributed congestion control algorithm is presented where now the forwarding decision rules are up to the source nodes. That is, the sources nodes will decide whether or not the packets are transmitted to their intended destination based on the current network state. For this, the relay nodes are in charge of notifying the changes to the source nodes according to the current network load. Finally, a machine learning model that predicts if a packet will be corrected transmitted or not to its destination based on the current network load is presented, implemented and evaluated from scratch.

- **Fairness between different sources.** Different solutions for the correct distribution of the network resources are evaluated. For this, the geographical position and the source rates of the nodes are taking into account, trying to avoid the monopolization of the network resources by some nodes.

All these objectives are described more extensively in the next section.

1.3 Summary of contributions

The main contributions of this thesis are described below:

- Modifications on the basic mechanisms and protocols used by the IEEE 802.11 mesh networks are proposed in order to improve their performance when using this technology as the implementation of the Smart Grid NANs. Mainly, a new multi-path mechanism is proposed and implemented in conjunction with a multi-channel allocation of the different available paths. These paths are assigned differentiately according to the quality of service demanded by every traffic. With this strategy, the proposed mechanism intends to take a better advantage of the available network resources, guaranteeing an adequate service to as much traffic as possible.
- Most Smart Grid applications have strong security and reliability requirements where their offered service can be reduced in both congestion and emergency situations. To this end, a congestion control mechanism, which takes into account possible emergency situations in the network, and applies also multi-channel allocation and traffic differentiation techniques, is presented. For this purpose, the default IEEE 802.11 mesh module has been again modified.
- A distributed congestion control mechanism that allows improving the performance offered by the NANs, when the selected technologies are now the IEEE 802.11 Wireless Ad Hoc Network (WANET) and IEEE 802.11ac physical layer standard is also proposed. The solution is conceptually simple as well as easily tunable and implementable. In addition, it is agnostic to the routing protocol, and of the MAC layer. The proposed mechanism combines several algorithms that allow differentiating the quality of service offered to each traffic based on its criticality, while providing a fair service to all nodes in the network.
- A machine learning-based congestion control mechanism is presented. This mechanism involves different task such as data collection, data preprocessing, training and validation of the trained model. This contribution is aimed to detail the most relevant aspects to implement a congestion control mechanism based on prior learning together with traffic differentiation. The proposed mechanism, based on the current network load, will decide whether or not to transmit a packet from the smart meters to the data concentrator.

The study and comparison of the benefits obtained is based on simulations carried out with the ns-3 network simulator [2]. Thanks to the improvements obtained, especially in situations of high load, NAN networks could offer service to a greater volume of traffic, and also allow a correct and differentiated quality of service for each application. In this way, new development challenges are created for both device manufacturers and electricity supplier companies, which will be able to offer new and better services to their customers. From our point of view, these new services should have a major impact on a greater efficiency in the use of electricity, and a faster and improved reaction in front of emergency situations.

1.4 Resulting publications

Most of the contents of this dissertation have been published in the following journals and conferences:

JCR Journal Publications:

- J. P. Astudillo León and L. J. De la Cruz Llopis, "A joint multi-path and multi-channel protocol for traffic routing in smart grid neighborhood area networks," Sensors, vol. 18, no. 11, 2018 [50].
- J. P. Astudillo León and L. J. De la Cruz Llopis, "Emergency aware congestion control for smart grid neighborhood area networks", Ad Hoc Networks, 93:101898, 2019. ISSN 1570- 8705 [51].
- J. P. Astudillo León, T. Begin, A. Busson and L. J. De la Cruz Llopis, "A Fair and Distributed Congestion Control Mechanism for Smart Grid Neighborhood Area Networks", Ad Hoc Networks, Ad Hoc Networks, page 102169, 2020. ISSN 1570-8705. [52].

Refereed Conferences:

- J. P. Astudillo Leon and L. J. de la Cruz Llopis, "Multi channel allocation and congestion control for smart grid neighborhood area networks," in Proceedings of the 15th ACM International Symposium on Performance Evaluation of Wireless Ad Hoc, Sensor, and Ubiquitous Networks, PE-WASUN'18. New York, NY, USA: ACM, 2018, pp. 1–8 [53].
- J. P. Astudillo León, T. Begin, A. Busson, and L. J. de la Cruz Llopis. "Towards a distributed congestion control mechanism for smart grid neighborhood area networks." In Proceedings of the 16th ACM International Symposium on Performance Evaluation of Wireless Ad Hoc, Sensor, and Ubiquitous Networks, pages 29–36, 2019 [54].

1.5 Outline of this Thesis

This thesis is organized as follows: Chapter 2 presents the background for the network technologies used. The next chapters presents the contributions of the thesis, and each of one includes the following: state of the art, description of the problem, proposed solution and numerical results. A more detailed explanation is presented below:

Chapter 3 presents a new proposal for improving performance in Smart Grid NANs when using IEEE 802.11 mesh network technology. Although the modification of the routing metrics is a good idea to differentiate the service offered to different traffics in the network, in this work the default routing metric is maintained and the work is focused on the modification of the mechanism used by the Hybrid Wireless Mesh Protocol (HWMP) protocol for the selection of the most appropriate path each time a data packet must be (re)transmitted. In this way, a modification of the HMWP protocol [55] is proposed and implemented, to allow an efficient selection of paths among multiple possibilities, depending on the service quality needs of the different traffic flows. The proposed mechanism is complemented with the assignment of different frequency channels to each available path.

Chapter 4 presents two different congestion control mechanisms for wireless multi hop wireless networks. Section 4.3 presents, a multi-channel allocation and congestion control mechanisms for Smart Grid NANs, when the different service quality needs of the different applications are considered. The congestion control mechanism takes into account if the network is in a state of emergency. For this, three network emergency states have been defined: normal, medium or high. Each of these states can be activated manually or automatically. For automatic operation, the network nodes periodically measure different performance parameters, and alert the rest of the nodes (by means of special management frames) in case of detecting anomalous situations. On the other hand, Section 4.4 presents a different solution to improve the network performance for NANs. Instead of considering IEEE 802.11s mesh networks, the IEEE 802.11 technology in ad hoc mode has been chosen where the recent IEEE 802.11ac standard was used. On the other hand, unlike the strategy presented in Chapter 3 and Section 4.3 where relay nodes are in charge of taking the forwarding decisions in network congestion situations, in this case the forwarding decision rules are up to the source nodes. This way, unnecessary transmissions of packets that will be likely discarded later on their way to their destination are avoided. Furthermore, the proposed mechanism is also designed to provide a fair distribution of the available network resources between all the source nodes, avoiding a higher utilization by the nodes with higher packet generation rate or simply favorably located. Furthermore, the solution is independent of the routing protocol used.

Chapter 5 presents a machine learning-based congestion control mechanism to improve the network performance for wireless multi hop networks. In this chapter, the different stages to generate the machine learning model are explained. In which it will be emphasized the importance of having a significant and meaningful dataset to train a model with a high predictive power (high accuracy to unseen data). In the first part of the chapter, the techniques for obtaining the samples and organizing them in a structured way are explained. In addition, mechanisms to select the most relevant features are detailed. Finally, the selection of the classifier, its training and validation are presented.

Finally, Chapter 6 summarizes the main conclusions of this dissertation and presents future lines of work where a comparison among the different proposed solutions are made.

Chapter 2

Background on Wireless Mesh Networks

This chapter makes a general review of the main features of the network technologies and protocols considered in this thesis. In this case, the wireless ad hoc network (WANET), and specifically, the standardized IEEE 802.11 mesh networking technology and their most prevalent protocols are covered.

2.1 Wireless Mesh Networks (WMNs)

Wireless ad hoc networks (WANETs) are a type of decentralized network since they do not need an existing infrastructure such as routers or access points. Therefore, each station (STA) participates in routing, that is, it retransmits packets from other nodes following rules calculated by a routing algorithm. It should be noted that routing decisions can change due to many factors in a dynamic network topology. As a particular case, WMN have attracted the attention of many researchers, among other characteristics, to the fact of not needing a central entity.

In the context of the wireless local area networks (WLANs) standardized by the IEEE, the proposal for multi-hop mesh networks was published in 2011 as the amendment number 10 to the 2007 general standard, with the name of IEEE 802.11s [56]. In the revision of the general standard published in 2012, as well as in the current revision [55], mesh networks have been directly incorporated, although a large number of researchers continue to refer to them as IEEE 802.11s mesh networks. The main differentiating characteristic of this type of networks is that, from the upper layers point of view, all the stations appear as connected at the MAC level although they might not be within the range of coverage. To make this possible, a layer 2 path search mechanism called Hybrid Wireless Mesh Protocol (HWMP), was designed.

Figure 2.1 shows the Wireless Mesh Network architecture, which consists of the following devices:

- **Mesh Station (MSTA):** 802.11 entity that supports mesh services. Mesh STAs participate in the formation and operation of the mesh. That is, in a multi-hop routing scheme, these entities can be source, destination or relay nodes.
- **Mesh Portal:** is a bridge or gateway device that allows the interconnection with external networks that are not IEEE 802.11.

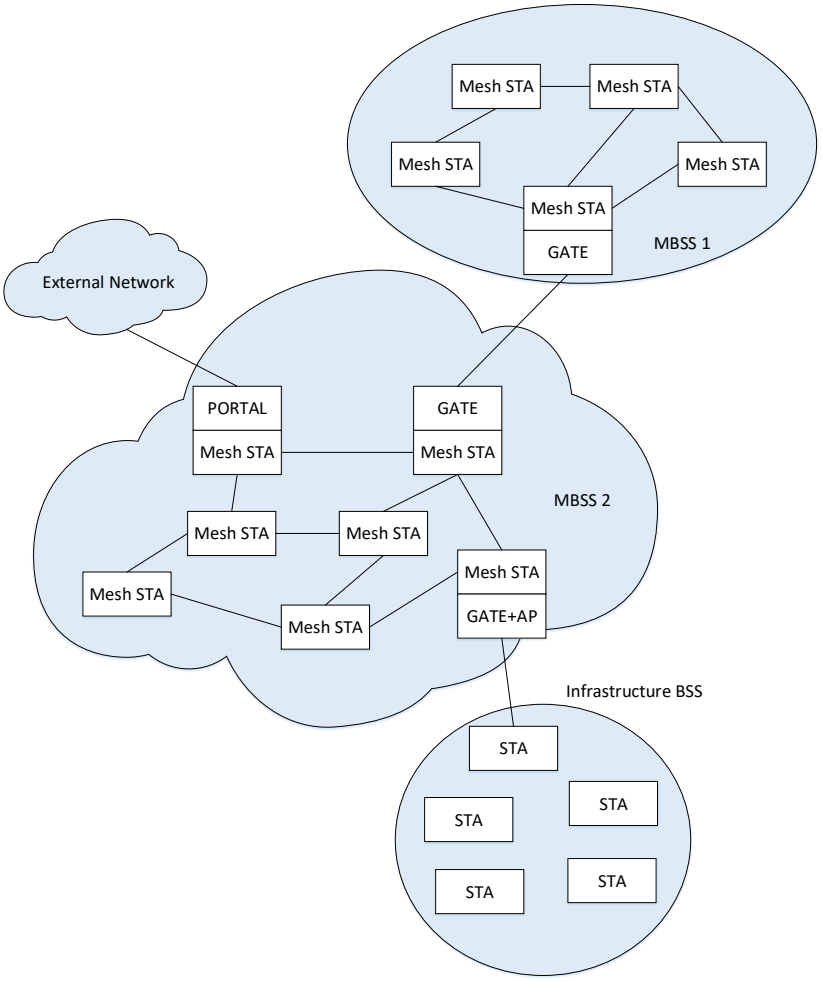


FIGURE 2.1: Wireless Mesh Network devices [24].

- **Mesh Gate:** allows interconnection between different mesh.

Very briefly, the steps that must be followed by the mesh stations to be able to provide their service to the higher layers are the following [55]:

2.1.1 Mesh Announce / Discovery

Mesh stations (MSTA) announce their presence and their availability to be part of a mesh network through special information in the beacon frames. All MSTAs can passively scan for beacon frames, or actively transmit Probe Request frames, in order to obtain the mesh profile, which is a set of parameters that contains the Mesh Basic Service Set (MBSS) configuration (mesh ID, path selection protocol, path selection metric, congestion control, synchronization method and authentication protocol).

2.1.2 Mesh Peering Management (MPM) protocol

When two stations detect each other, they establish a link by means of the exchange of two action frames (Peering Open and Peering Confirm). These links are periodically maintained, and therefore every station has an updated knowledge of its available neighbors. This protocol maintain bidirectional links, that is, when two potential neighbors have successfully transmitted and received Mesh Peering Open and Confirm frames (2.2) [24].

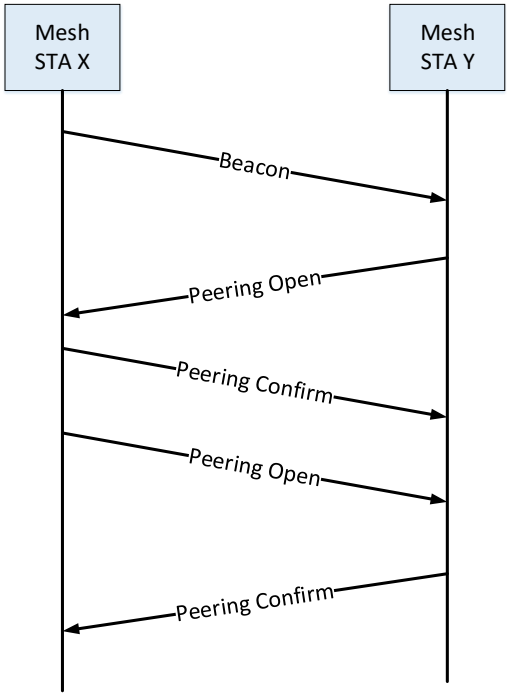


FIGURE 2.2: Mesh Peering link stablishment handshake.

2.1.3 Medium Access

It combines the contention-based scheme of Enhanced Distributed Channel Access (EDCA) with a contention-free scheme called MCF Coordinated Channel Access (MCCA) [7]. These mechanisms are described as follows:

- **Enhanced Distributed Channel Access (EDCA):** EDCA uses four access categories (ACs) to provide traffic differentiation: voice (Vo), video (Vi), best effort (BE), and background (BK). Each category has its own set of medium access parameters: the arbitrary interframe space number (AIFSN), the contention window minimum and maximum values (CWMIN and CWMAX), and the optional transmission opportunity limit (TXOPLimit) [7].

  When EDCA is not used, the default medium access in IEEE 802.11 is the Distributed Coordination Function (DCF), which is devoted to use only for best effort services. That is, all traffic types compete with the same opportunities to access the channel. However, for real-time applications it is mandatory to ensure minimum quality of requirements such as bandwidth, delay and delay jitter. In order to meet these requirements, 802.11e proposes EDCA.

- **MCF Coordinated Channel Access (MCCA)**. MCCA provides contention-free transmission using a resource reservation mechanism. Each reservation consists of a set of time intervals refered as MCCA transmission opportunities (MCCAOPs). With this strategy, the MCCAOP owner (the station that performed the reservation) may transmit to the MCCAOP responders (the stations that receive the reservation request) [7].

2.1.4 Path Mesh Selection and Forwarding

To send packets between a specific origin and destination, a path must be discovered, which will generally expand through multiple intermediate hops. This path is built by means of the HWMP protocol, which combines the flexibility of on-demand path selection with proactive tree topology extensions. It enables efficient path selection with or without access to the infrastructure, and supports two types of path selection modes:

- **On-demand mode**: This mode is always available, independently whether a root MSTA has been configured or not. When a new data frame arrives from the application or physical layer, it is immediately stored at the HWMP queue. If the frame must be forwarded, the mesh STA selects the best path for its transmission. For this purpose, the HWMP checks the destination address and looks up the next hop address in the routing table, previously calculated by the path-building mechanisms. If a path is available, the mesh STA transmits the frame by using the medium access mechanism. On the other hand, if there is not a route to the destination, the path discovery mechanism is activated (Figure 2.3). For this purpose, a Path Request (PREQ) message is broadcast from the source to the whole network, and when this message reaches the intended destination, this node replies with an unicast Path Reply (PREP) message to the PREQ originator node by using the best path (lower metric) [57]. Finally, in the case of one or more unreachable destination(s), the mesh STA drops the frame and transmits control messages such as path error (PERR) in order to invalidate the whole path. This message is sent to all traffic sources that have a known active path to the destination(s). The active forwarding information associated with the unreachable destination(s) should no longer be used for forwarding [55].

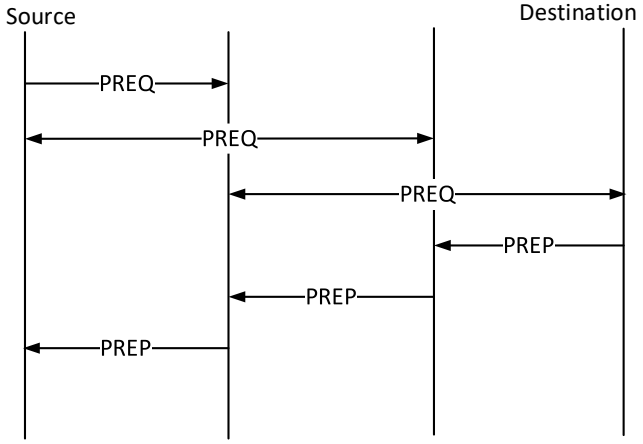


FIGURE 2.3: On-demand mode.

- **Proactive mode**: When one of the network nodes can be considered as the main source or destination of the data packets (for example, a gateway that provides Internet access),

it is more convenient to use the proactive mode (Figures 2.4 and 2.5). In this mode, at least a MSTA must be configured as root MSTA, and additional proactive tree building functionality is added to the on-demand mode. The root MSTA has two options to start the path search mechanism: sending proactive PREQs messages to the rest of MSTAs by using group addressed communication, or broadcasting root announcement (RANN) messages, which are then used to initiate a path setup by the rest of MSTAs using acknowledged communication.

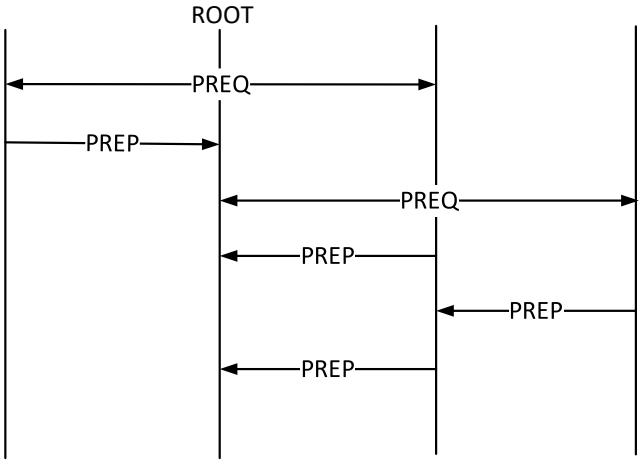


FIGURE 2.4: Proactive PREQ-based.

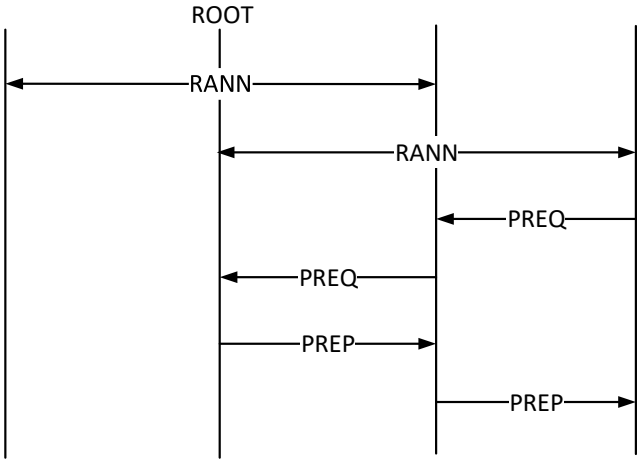


FIGURE 2.5: Proactive RANN-based.

2.1.5 Link metric

HWMP uses by default the airtime link metric for path selection:

$$C_a = \left[O + \frac{B_t}{r} \right] \frac{1}{1 - e_f} \tag{2.1}$$

where O is the channel access overhead; B_t, the transmission frame size; r, the data rate; and e_f , the error rate. This metric indicates the amount of time needed to transmit a frame over a link.

2.2 Routing Protocols

There are several routing protocols for ad hoc and mesh networks, and they are evaluated and analyzed continuously in the scientific community [58]. Several surveys [59, 60] provide extensive descriptions of the different routing protocols. Some of the most relevant routing protocols for WMNs are: Ad hoc On-demand Distance Vector (AODV) [61], Optimized Link State Routing (OLSR) [62, 63], BATMAN (Better Approach To Mobile Adhoc Networking) [64] and BMX6 [65, 66], and they are briefly described below.

2.2.1 Ad hoc On-demand Distance Vector (AODV)

The AODV [61] protocol is very similar to the previously explained HWMP. In fact, HWMP is based on ADOV. In AODV, the path discovery mechanism is based on the Route Request (RREQ) and Route Reply (RREP) messages, which have the same functionality as PREQ and PREP frames in HWMP. When a node requires a route to the destination, a RREQ message is broadcasted to the whole network. When the RREQ reaches the intended destination, it replies with a unicast RREQ to the source node building the path. Finally, the neighboring nodes are maintained through broadcasting HELLO messages. As with AODV there is no a peering protocol, the network nodes must discover their neighbours. To this end, special HELLO messages are periodically transmitted by every node to announce itself to the rest of the network nodes.

2.2.2 Optimized Link State Routing (OLSR)

The Optimized Link State Routing protocol is defined as an optimization of the classic link state algorithm but focused on mobile networks. OLSR is suitable for networks where traffic is random and sporadic, and there it not assumptions about the underlying link layer. The protocol does not require reliable transmission of control messages, since each node periodically sends such messages in certain time intervals. In addition, each control message contains a Sequence Number (SN) that is incremented for each message. SNs allow to identify at the reception which information is most recent. There are currently two versions of this protocol: OLSRv1 (IETF RFC 3626) [62] and OLSRv2 (IETF RFC 7181) [63].

Protocol Features

MultipointRelay nodes (MPR) are the key concept of OLSR[1]. These nodes optimize the classic link state algorithm in three ways:

- MPRs are a set of nodes that forward broadcast messages during the flooding process. With this strategy, the overhead of control messages are reduced as compared to a classic flooding mechanism (Figure 2.6).

[1] It is demonstrated that the use of MPRs improves the network scalability [67]

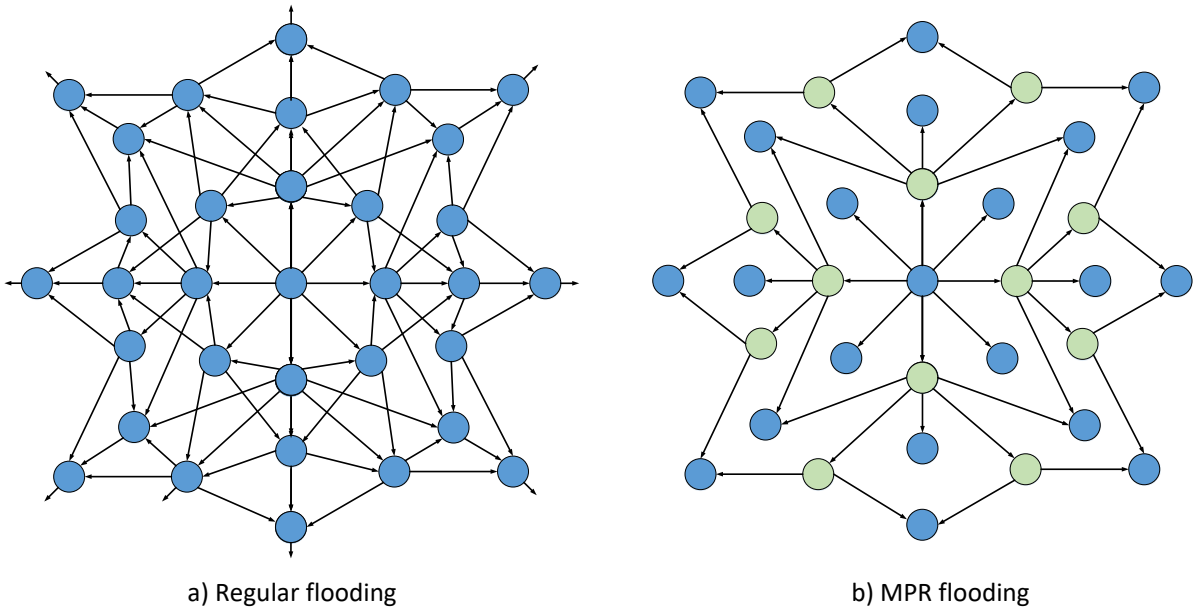


FIGURE 2.6: Regular flooding vs MPR flooding [68].

- Link state information is generated by MPR nodes. It minimizes the number of control messages flooded in the network.
- An MPR node distributes the link state information in the network.

Optimized Link State Routing version 2 (OLSRv2)

OLSRv2 [63] maintains the basic mechanisms and algorithms of OLSRv1 where additive and bidirectional link metrics are now included[2]. OLSRv2 adds new message types and modifies the existing messages. The Neighborhood Discovery mechanism is performed by Neighborhood Discovery Protocol (NHDP)[69].

Each node in OLSRv2 must select two MPR sets:

- **Flooding MPRs:** These nodes are used for flood reduction. The MPR flooding operation is used when control messages are broadcast to the whole network [63].
- **Routing MPRs:** These nodes are used to reduce the topology [63].

OLSRv2 uses metric-based routing, that is, it allows the links to have a metric chosen. With this routing protocol, the metrics are additive and the routes are created taking into account the the minimum sum of the link metrics along that route [63, 70].

[2]OLSRv1 uses the hop count metric

Multi-Topology Extension for OLSRv2

Multi-Topology extension allows a router to establish and maintain multiple routing topologies. Each topology is associated with a link metric type. Routers can belong to one or more topologies. That is, each router maintains a routing set for one or more topologies, allowing separate packet routing for each topology [71].

2.2.3 Better Approach To Mobile Adhoc Networking (BATMAN)

BATMAN [64] is a network layer routing protocol where each participating node just learns the best next neighbor for each destination. For routing decisions, this protocol uses the lost packets due to unreliable links. BATMAN handles the concept of collective intelligence. That is, the topology information does not rely by a single node. BATMAN uses three important concepts:

- **Originator Message (OGM):** This message announces the existence of an Originator. OGMs are used to determine the link quality and path selection.
- **Originator:** It is a network interface which is announced by the Originator Messages.
- **Best Link:** It is the better interface or better next hop node to a given node Originator.

Sequence Numbers are the key information that is transmitted with each OGM. Sequence Numbers are recorded in Sliding Windows. The Sliding Windows always contains the set of recent received Sequence Numbers. The amount of Sequence Numbers recorded in the Sliding Window allow to calculate the metric.

BatMan-eXperimental version 6 (BMX6)

BatMan-eXperimental version 6 [65] is the successor of BMX (BMXd) which emerged as an independent branch of BATMAN. This protocol is focused on IPv6 addresses where simplified message dissemination is used. BMX6 manages to reduce network overhead by two different mechanisms. First, it optimizes the traffic transmitted through the network using compact IID and hashes description (neighbors). Second, flooding reduction is achieved by analyzing whether a link is relevant or not [65].

Chapter 3

Multi-Path and Multi-Channel Routing

Among the possible mechanisms to improve the performance of NAN networks, in this chapter the work is devoted to the path discovery mechanism when the selected technology is the wireless mesh networks. To this end, some general modifications are proposed for the routing protocol of the wireless multi-hop mesh networks standardized by the IEEE. In particular, the possibility of using multiple paths and transmission channels at the same time, depending on the quality of service needs of the different network traffic, is added.

3.1 Introduction

In order to improve the management mechanisms of the electric energy transport infrastructures, the smart grid networks have associated data networks that are responsible for transporting the necessary information between the different elements of the electricity network and the control center. Besides, they make possible a more efficient use of this type of energy. One of these data networks, the Neighborhood Area Network (NAN) is responsible for interconnecting the different smart meters and other possible devices present at the consumers' premises with the control center. Among the proposed network technologies for NANs, wireless technologies are becoming more relevant due to their flexibility and increasing available bandwidth. In this chapter, the routing mechanism used by the IEEE 802.11 mesh networks is modified, when this technology used as the implementation of the Smart Grid NANs. Mainly, the path discovery mechanism is modified in order to obtain all possible paths between the smart meters and the data concentrator. The proposed solution ranks the paths based on the airtime link metric. With this strategy, priority traffic types are transmitted for the best paths, while the traffic types with lower QoS needs are transmitted for the worst ranked paths. Remember that airtime link metric is cumulative, and the best paths are those with lowest metric. In order to improve even more the benefits obtained, priority traffics are also transmitted for the less congested channels. For this reason, traffic differentiation at mac level is also proposed.

The proposed modifications have been implemented in the ns-3 simulator and evaluated in situations of high traffic load. Simulation results show improvements in the network performance in terms of packet delivery ratio, throughput and network transit time.

The rest of the chapter is organized as follows. Section 3.2 presents the related work. Next, Section 3.3 presents the modifications proposed for the HWMP protocol. Details about the multi-path and multi-channel mechanisms implementation are provided, as well as the route selection and assignment algorithms. Section 3.4 presents and analyzes the results obtained through the simulations and, finally, the conclusions as well as the future lines of research are summarized in section 3.5.

3.2 Related works

Smart Grid networks have attracted the attention of numerous researchers in recent years. Among these investigations, several proposals have been presented in order to improve the benefits offered by the NANs, where both wired and wireless technologies have been taken into account [72]. Within the wired technologies, PLC stands out especially in this environment due to its ability to use the existing infrastructures. However, the available bandwidth with this technology is quite limited and it also presents drawbacks when data signals must pass through electrical transformers. When the number of nodes in the network grows, as well as the bandwidth needs, other technologies must be considered. In this sense, wireless networks in general [73] [74] and wireless multi hop networks in particular [75] [76] present a series of advantages that make them ideal candidates. For instance, they do not require previous infrastructures and their bandwidths are constantly increasing. Besides, they have a great flexibility to modify the network topology and to take advantage of multi-channel and multi-path mechanisms that increase their performance in terms of, among others, availability, packet delivery ratio or network transit time. For these multi-hop wireless networks, a new and precise analytical model, which takes into account the hidden nodes problem, has been presented in [77].

In [78] an enhancement of the Optimized Link State Protocol (OLSR) in order to satisfy the required level of reliability in NANs is presented. The possibility of offering an adapted quality of service to the different data traffics transmitted through the network is taken into account. To this end, authors propose the use of a combination of different basic metrics: Expected Transmission Count (ETX), Minimum Delay (MD) and Minimum Loss (ML). They choose Relevant Link Metric Types (RLMTs) for each application (traffic type), assign different weights to each of them, and use a pruning technique to reduce the number of considered paths to a given destination. The best link to send each traffic is then calculated by means of an AHP (Analytical Hierarchy Process) algorithm. The proposal is evaluated by means of ns-2 simulations, over an usual network environment consisting on a grid of smart meters transmitting (receiving) information to (from) a data concentrator, and taking into account four basic CBR traffic types. Moreover, the topology is modified increasing the number of smart meters (from 25 to 64) and changing the data concentrator position. The network performance is measured in terms of dropped packets, packet delivery ratio and average delay, showing a better behavior when compared to a basic OLSR implementation. Same authors had previously presented in [79] a performance evaluation and comparison of OLSR and HWMP (IEEE 802.11s) routing protocols, together with a classification of the main AMI application traffics.

A multigate communication network, based on IEEE 802.11s, is proposed in [73] for Smart Grids. Authors take into account the possibility of having more than one node acting as a gateway, together with a real-time traffic scheduling and a multi-channel aided routing protocol. Besides, authors propose a heuristic backpressure scheme, where every node evaluates the state of its neighbors before selecting one of them as the best next hop, which implies that some information

(the backpressure metric) must be periodically exchanged between nodes. Otherwise, to avoid loop problems, a hop-count limit is imposed to the data packets. Besides, in order to reduce the effect of cochannel interference, a multi-channel protocol is also introduced. To evaluate the proposals, three simulation scenarios are taken into account: a) three separated sub-networks where every one has its own gateway, b) a multigateway network where the three previous networks share their three gateways and where the nodes are uniformly distributed, and c) the previous configuration but with an asymmetrical distribution of the nodes. The results show the better behavior offered by the proposed backpressure scheme in terms of overall throughput, average end-to-end delay and adaptation to malfunctioning nodes. On the other hand, the benefits of the multi-channel protocol are also clearly shown.

A cross-layer mechanism which combines information from the physical, MAC and network layers is presented in [80]. Based on that mechanism, authors define a new routing metric (Expected Path Throughput, EPT) and a distributed routing protocol, which is evaluated with the help of the ns-2 simulator. The results show the good behavior of the proposal when compared with other classical metrics and protocols.

In [81], authors propose the HWMP-NQ protocol, a modification of HWMP to ensure the needed quality of service (QoS) of several smart grid traffic types. To this end, the airtime link metric is modified by considering the packet size and the transmission rate. However, the needed number of channel measurement could be excessively increased. To avoid this, a frame error rate computing algorithm based on a single measurement is also proposed. Besides, the benefits provided by a multi-gateway backup routing scheme are also analyzed. Moreover, to reduce the routing overhead in case of link failures, a modification of the path error mechanism is introduced. To evaluate the benefits of their proposals, authors build classical NAN grid topologies with the help of the ns-3 simulator, and run multiple simulations to measure the average throughput, packet delivery ratio, end to end delay and routing control information overhead. The results show the benefits of the multigate routing scheme presented in [73] and the HWMP-NQ protocol, in front of the classical HWMP implementation, for different NAN grid sizes (from 9 to 64 smart meters). What is more, the influence of the nodes failure rate is also studied, showing that the performance improvements obtained with the authors proposals increase when that failure rate is higher.

In order to improve the network throughput and reliability, another modification of the airtime link metric calculation method was presented in [75]. One of the contributions of this work is to give more importance to the upstream transmission status (from smart meters to the concentrator), since most data is transmitted in this direction. Besides, a modification of the path selection mechanism is provided to avoid the classical problem of route fluctuation. With this modification, not only the current airtime link metric value, but also its variations, are taken into account to select (or not) a new route between two network nodes. Ns-3 simulations are presented to show the achieved benefits in terms of packet delivery ratio, end-to-end delay and data retransmission count. The results also highlight the need for congestion control mechanisms when the network size is increased.

Some of the same authors of [75] make in [76] a study of the HWMP routing protocol, with the goal of identifying its weakness, both from the HWMP protocol itself (route instability and route recovery) and from the integration with Smart Grid networks (oversimplified calculation of airtime link metric and the need of traffic differentiation). Here, a modification of the airtime link metric computation is also proposed, as well as a proposal for the path selection mechanism. Besides, to get a better performance in terms of packet losses, reserve routes are stored in the network nodes. This idea gives rise also to a reduction in the traffic management traffic

needed when a path is broken. Moreover, in order to provide a better quality of service to some applications, a delay-tolerant traffic management method based on the concept of delay-tolerant networking is proposed. The improvements obtained with the application of these new solutions to the protocol (called HWMP-RE) are checked and shown by means of ns-3 simulations. Grid topologies are considered, from 9 to 49 nodes, where every node generates traffic (belonging to seven different applications) to two root mesh stations (gateways). HWMP-RE is compared with the basic HWMP and with the previous proposal in [75], showing a better behavior in terms of packet delivery ratio, end-to-end delay, number of PERR/PREQ generations, throughput and reliability.

Other proposals based on the modification of the HWMP metric can be found in [82] and [83]. In [82] a QoS-aware and load-balance routing scheme is proposed, which is complemented with an EDCA based adaptive priority adjustment scheme, with the goal of satisfying the QoS requirements of different NAN applications. The modification proposed for the airtime link metric consist of including the packet size and calculating the frame error rate separately for the different NAN applications. Besides, to avoid congested paths, the queuing delay is also added to the metric. What is more, the dynamic adjust of the packet priority allows a better resources utilization under low load conditions, and improve the reliability under heavy load conditions. Ns-3 simulations are carried out to evaluate the obtained performance, which shows an increase of both the packet delivery ratio and the throughput, as well as a reduction of the average end-to-end delay. The network scenario consist of a grid topology where the number of nodes varies between 9 and 64.

On the other hand, the metric modification proposed in [83] (interference aware expected transmission time, IAETT) is oriented to reduce the impact caused by inter and intra-flow electromagnetic interferences. Besides, traffic differentiation is also considered. Based in this metric, an interference aware QoS routing protocol is proposed and evaluated. The performance evaluation is carried out again by ns-3 simulations, over a scenario consisting of 100 nodes arranged in a 10x10 regular grid, where both the gateways (nine nodes) and the traffic generating nodes are randomly chosen. Results show the improvements obtained in terms of average end-to-end delay and packet delivery ratio.

As already mentioned, in this chapter a new proposal for improving performance in Smart Grid NANs when using IEEE 802.11 mesh network technology is presented. Although the modification of the routing metrics is a good idea to differentiate the service offered to different traffics in the network, it has been preferred to maintain the basic airtime link metric and focus the efforts on the modification of the mechanism used by the HWMP protocol for the selection of the most appropriate path each time a data packet must be (re)transmitted. By its own nature, the default metric informs about the congestion state of the different network areas, which is the most relevant measure for the present approach. Moreover, it is important to keep in mind that working with more complicated metrics usually lead to higher CPU and memory requirements in the network nodes, as well as to protocols that generate more network control traffic. In this way, a modification of the HMWP protocol is proposed and implemented, to allow an efficient selection of paths among multiple possibilities, depending on the service quality needs of the different traffic flows. The proposed mechanism is complemented with the assignment of different frequency channels to each available path. In addition, to avoid packet losses due to the formation of unwanted loops, the proposed technique is combined with a criterion of minimum number of hops when choosing the paths. This technique reduces the number of selectable paths, but avoids the need of using packet hop counters (which are used by the nodes to discard packets after a given number of hops, with the added disadvantage of using network resources for a certain number of retransmissions in a completely useless way). On the

other hand, as will be seen in the results section, it has been considered of great importance to provide not only the average values of the performance parameters under study, since this way the real network performance is not obtained and would probably lead us to an erroneous network planification.

In the next section, the proposed solution is explained.

3.3 Proposed solution

Figure 3.1 shows the proposed modified structure for the HWMP algorithm. On the one hand, the MSTAs are capable of storing multiple paths to every destination node in their routing tables. These paths are classified by a path selection policy with the objective of sending the data traffic with the highest priority over the best paths. On the other hand, to reduce the level of interference between MSTAs and increase the network performance, a different channel is assigned to each available path. Besides, a different channel will be reserved for control packets. In order to add these multi-path and multi-channel functionalities to the default HWMP protocol, several mechanisms are proposed. To evaluate their performance, all the proposals have been programmed and included in the basic ns-3 IEEE 802.11s module [2].

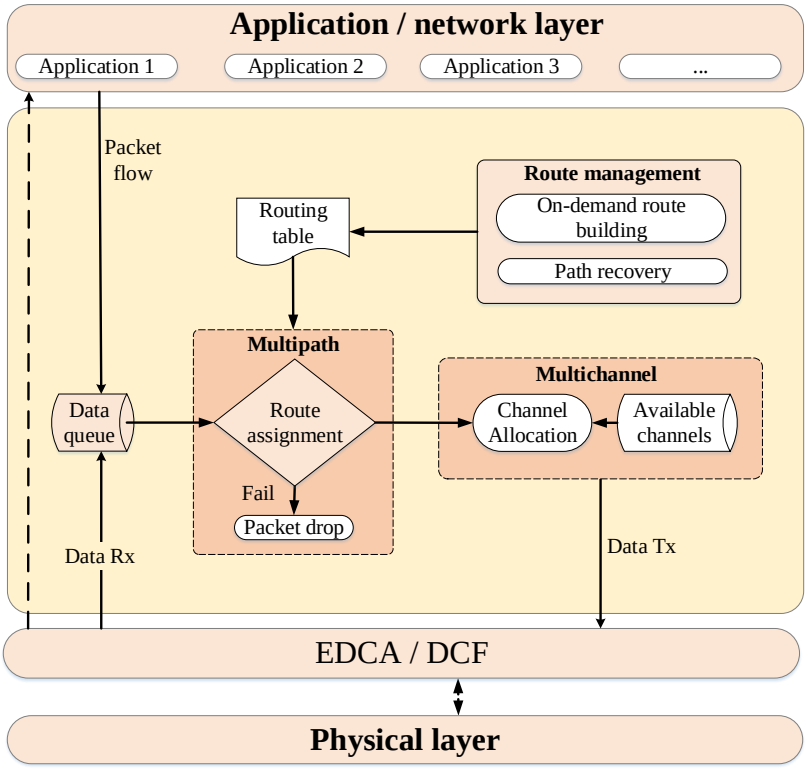


FIGURE 3.1: General view of the multi-path and multi-channel modules inclusion in HWMP.

3.3.1 Multipath proposal and implementation

HWMP establishes by default a single path between the source and destination nodes. The purpose of this subsection is to modify the protocol to obtain and take into account all the

possible paths between two nodes. As already said, these available paths will be assigned to the different applications (traffic types) depending on their priority. Therefore, the modules related to the route management, routing table and route assignment have been modified.

Route management

To allow the existence of more than one path between each pair of nodes, it is first necessary to modify the acceptance criteria for both PREQ and PREP packets. Table 3.1 summarizes all the functions and variables needed for the new PREQ-PREP mechanism. As shown in the Algorithm 1, first of all the most relevant fields are extracted. Then, the metric value for every path is updated. Next, the algorithm has to validate the message, that is, verify if the current message has more recent information (the sequence number of the current message is greater than the previous one, $SN_{C_{OA}} > SN_{P_{OA}}$) or if there is a better metric when the sequence number is the same ($SN_{C_{OA}} = SN_{P_{OA}}$). By default, HWMP updates the route to the originator address and replaces the previous route when the sequence numbers of multiple received PREQ messages are equal but one of them has better metric. For instance, in the example shown in Figure 3.2(a), the source node S generates a PREQ message to find a path to the destination node D. Two instances of this PREQ (first $PREQ_1$ and then $PREQ_2$) are received by N_3 from two different nodes (paths), and only the one with the best metric will be retransmitted to D ($PREQ_1$ in the figure). In this proposal, the $PREQ_2$ message is also retransmitted (Figure 3.2(b)), because there is the need to compare not only the sequence numbers and the metrics in order to validate a PREQ message, but also take into account the previous (retransmitter) node in the path (f_r in Table 3.1 and Algorithm 1). The objective is to maintain multiple paths to the originator address (OA) through different f_r nodes. After validating the message, the routing table is updated (specifically, the table entries related to the neighbor and to the source nodes). In addition, if there were queued packets for the new or updated route, they would be immediately transmitted. In the case that the PREQ destination address is the one of the receiving node, it means that a route has been found. Therefore, this node transmits directly a unicast PREP message towards the S node.

In the same way, with the default algorithm the destination node responds with a PREP message to the source node if and only if the PREQ received has better metric than the previous ones (Figure 3.3(a)). However, in this implementation a PREP message is sent although the received PREQ has worse metric. This action allows to propagate not only the best path to the source, but to have several paths with different metric values (Figure 3.3(b)).

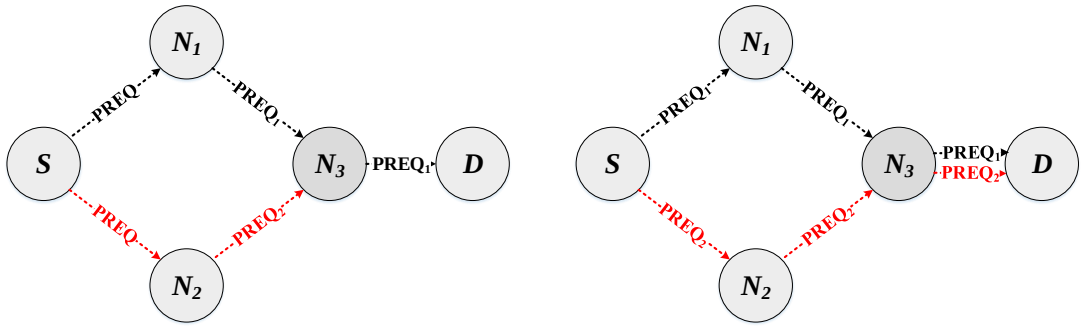
TABLE 3.1: Definition of the variables and functions for the PREQ and PREP mechanisms.

Parameter	Description
preq or **prep**	PREQ or PREP element received in the node.
read	Function that extracts the main fields of the **preq** or **prep** elements.
$\mathbf{OA_C}$	Originator address of the current element received.
$\mathbf{OA_P}$	Originator address of the previous element received.
$\mathbf{SN_{C_{OA}}}$	Sequence number of the current element received to the $\mathbf{OA_C}$.
$\mathbf{SN_{P_{OA}}}$	Sequence number of the previous element received to the $\mathbf{OA_P}$.
$\mathbf{DA_C}$	Destination address of the element.
$\mathbf{m_{1\cdots n}}$	Cumulative metrics of each path.
$\mathbf{f_r}$	Retransmitter node.
pathID	Path identifier
updateMetric	Calculate and update the ALM metric according to the received **preq** or **prep** message.
isValid	Verify if the element received has updated information (sequence numbers), or it has better metric or different f_r
updateTable	Update or add a new route if the **isValid** function is true.
RouteSelection	Request a route to the destination (next hop), taking into account the application.
createPrep	Create the PREP message according to the parameters of the PREQ message.
sendPrep	Send the PREP unicast message to the $\mathbf{OA_c}$
$\mathbf{forward_{control}}$	Retransmits the **preq** broadcast message through all interfaces.

Algorithm 1: PREQ elements process and forwarding.

Input: preq message
Output: Update the routing table

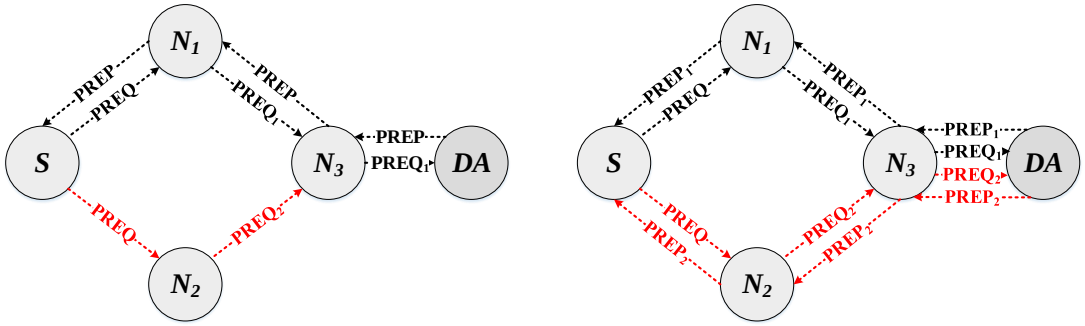
```
receive_control(preq)
[SN_c, DA_c, OA_c, m_1...n, pathId, f_r] ← read(preq)
preq ← updateMetric(m_1...n)
if isValid(preq) then
    updateTable(SN_c, OA_c, m_1...n, f_r, b_c)
    packets ← getPacketsInQueue()
    while packets == EMPTY do
        p_q ← getPacket(packets)
        [next_hop, isThereRoute] ← RouteSelection(p_q, DA_c)
        channel_data ← channelAllocation(p_q)
        send(p_q, next_hop, channel_data)
else
    return
if DA_c = myA then
    createPrep() ⟵ preq
    sendPrep(prep)
else
    forward_control(preq, channel_control)
```



(a) **HWMP.** N_3 receive, process and forward the $PREQ_2$ message if and only if it has better metric than $PREQ_1$.

(b) **MPC-HWMP.** N_3 receive, process and forward the $PREQ_2$ to maintain multiple paths.

FIGURE 3.2: Modification to the acceptance criteria when the intermediary nodes receives a PREQ message.



(a) **HWMP.** The destination node replies with an unicast PREP message to the source node.

(b) **MPC-HWMP.** The destination node replies all received PREQ with unicast PREP messages to the source node through its differents paths.

FIGURE 3.3: Modification to the acceptance criteria when the destination node receives node a PREQ message.

The main modifications for the reception, processing and forwarding of a PREP message are similar to the explained for the PREQ messages, and are detailed in the Algorithm 2. On the other hand, the criteria for the retransmission of the PREP messages towards the node that originated the PREQ message, must consider the multiple paths created and not take erroneous paths. In other words, each PREP message must know which was the path that took its corresponding PREQ message. To this end, a field has been added both to the PREQ-PREP messages and to the routing table ($pathId$), which allows the nodes to obtain the correct path (destination address) for each PREP message that must be generated or forwarded. For instance, in the example shown in Figure 3.3, N_3 make use of this field to route correctly $PREQ_1$

and $PREQ_2$ to their corresponding nodes.

Algorithm 2: PREP elements process and forwarding.

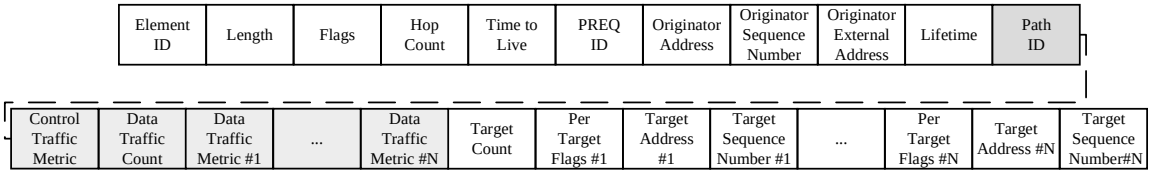
Input: prep message
Output: Update the routing table

```
1  receive_control(prep)
2  [SN_c, DA_c, OA_c, m_1...n, pathID, f_r] ← read(prep)
3  preq ← updateMetric(m_1...n)
4  if isValid(prep) then
5      updateTable(SN_c, OA_c, m_1...n, f_r, b_c)
6      packets ← getPacketsInQueue()
7      while packets == EMPTY do
8          p_q ← getPacket(packets)
9          [next_hop, isThereRoute] ← RouteSelection(p_q, DA_c)
10         channel_data ← channelAllocation(p_q
11         send(p_q, next_hop, channel_data)
12 else
13     return
14 if DA_c = myA then
15     routeResolved()
16 else
17     findRoute(prep, pathID)
18     forward_control(preq, channel_control)
```

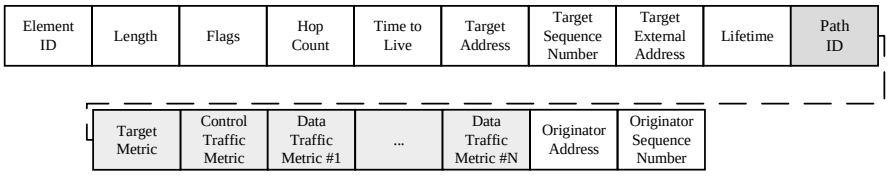
3.3.2 Multi-Channel mechanism

The objective of implementing multi-channel techniques is to increase the network performance mainly in stress situations, that is, when a high amount of data traffic is being transmitted by the network. This proposal implements traffic differentiation in the following way. First, different channels for control and data traffics have been used, where just one channel is assigned to the control traffic (route management). On the other hand, as explained in the previous subsection, the different data traffics are mapped to the available paths depending on their priority, and a different channel is assigned to every path. Thus, each channel will have a specific metric value which will be useful for future routing decisions.

The ALM metric is cumulative and updated in each node by the Path Request (PREQ) and Path Reply (PREP) messages. Therefore, to propagate the ALM metrics according to their respective channel it is necessary to make modifications to the structure of the PREQ and PREP messages, as can be seen in Figure 3.4. The modification consist in the inclusion of the path identifier field, the metric value for the control channel, the number of available paths (channels), and the metric value for each of them.



(a) Path Request.



(b) Path Reply.

FIGURE 3.4: Path Request and Path Reply modifications.

By default, the HWMP protocol performs a broadcast of PREQ messages for path discovery through all the available interfaces in the node. Therefore, with the multi-channel implementation, the number of broadcast messages could be increased excessively. This is the reason why a specific channel has been defined for control messages, thus avoiding a high and unnecessary load on the data channels.

To implement the path/channel allocation, the applications are marked from the source with an Enhanced Distributed Channel Access (EDCA) category [55]. EDCA distinguishes between four types of traffic according to their QoS needs. The types of traffic with higher priority are mapped to the highest categories of EDCA (Voice or Video), and vice versa. Therefore, intermediate nodes are able to select the next hop node from their routing table among the multiple available paths to the destination, and transmit the application traffic over the correct channel.

3.3.3 Routing selection and assignment

The proposed mechanisms modify the default HWMP routing table. On the one hand, the number of entries in the table will be higher, due to the availability of multiple paths for each destination address. In addition, the number of fields of each entry will also be higher to allow the appropriate path selections. The added fields are summarized in Table 3.2.

TABLE 3.2: New fields added to the HWMP routing table.

Field	Description
Control channel metric	Metric value of the control channel, obtained by the PREQ-PREP mechanism.
Data channel metric	Metric value of each data channel, obtained by the PREQ-PREP mechanism.
Hop count	Number of hops between two nodes.
PathId	Path identifier used by PREP messages.

The general route assignment tasks performed by the network nodes every time they have to (re)transmit a packet are detailed in the Algorithm 3. First, the node extracts the following parameters from the packet header: source node, destination node, access category and time to live. Then, the algorithm verifies if there is an available route. In the affirmative case, the next hop node is obtained by means of the route selection algorithm (Algorithm 4), considering the destination node and the access category, and finally the transmission is assigned to a specific channel. In the case that there is no route, the packet is queued and the path discovery mechanism is activated. This mechanism tries to obtain the route a fixed number of times and if that threshold is exceeded the route to that destination is considered invalid and the packet is eliminated.

Algorithm 3: Route assignment

Input: NAN application
Output: Forward the data to the next hop depending of the multiple paths available at the node.

```
1  receive_data(packet)
2  [source, destination, accessCategory, TTL] ← read(packet)
3  [next_hop, isThereRoute] ← RouteSelection(packet, destination)
4  if isThereRoute then
5      channel_data ← channelAllocation(packet)
6      send(packet, next_hop, channel_data)
7      return
8  else
9      if shouldInitiatePathDiscovery(destination) then
10         lastSQN ← getLastSQN(destination)
11         preq ← createPreq(lastSQN, destination)
12         sendPreq(preq)
13     queuedPacket(packet)
```

As it can be seen in Algorithm 4, the route selection process first searches all available paths according to the destination address and then delete the routes that have expired. Later, these paths are sorted from the best to the worst according to the metric, and then they are resorted taking into account the number of hops to te destination. This implementation considers the number of hops to avoid the creation of undesired loops, as will be explained with the help of Figure 3.5. This figure represents a simple scenario with four nodes: source (S), destination (D) and two intermediate nodes (N_1 and N_2). The source node has two available paths to send its packets to the destination, $P1_{S\to D}$ through N_1 and $P2_{S\to D}$ through N_2. Suppose that at a given moment the metric value of $P1_{S\to D}$ is better than the metric value of $P2_{S\to D}$. Therefore, according to the multi-path mechanism, high priority packets will be sent to N_1 and low priority packets will be sent to N_2. Similarly, N_2 has two available paths to D, one with better metric value directly to D, $P1_{N_2\to D}$, and another with worse metric value $P2_{N_2\to D}$ through S. In this way, low priority packets would be sent back to S, building a loop from which they would never leave. To avoid this problem, the criteria of minimum number of hops is also taken into account, so that N_2 will never use the $P2_{N_2\to D}$ path, sending all the packets with destination in D directly to D regardless of their priority. This way, the source node S is allowed to use its two available paths, sending high priority packets through N_1 and low priority packets through N_2, but N_2 must use always the same path to D. As previously said, this mechanism reduces the number of available paths, but avoids the creation of undesired loops, eliminating the need for packet hop counters and unnecessary retransmissions which consume network resources in a completely useless way. Finally, the algorithm maps applications with the highest priorities to

the best paths.

Algorithm 4: Route Selection

Input: packet, destination
Output: Get the next hop address

```
accessCategory ← getEDCA(packet)
nextHops ← getNextHops(destination)
nextHops ← eraseExpiredRoutes(nextHops)
if len(nextHops > 0) then
    nextHops ← sortByHopCount(nextHops)
    nextHops ← sortByMetric(nextHops)
    next_hop ← getPath(nextHops, accessCategory)
    return[next_hop, True]
else
    return[any, False]
```

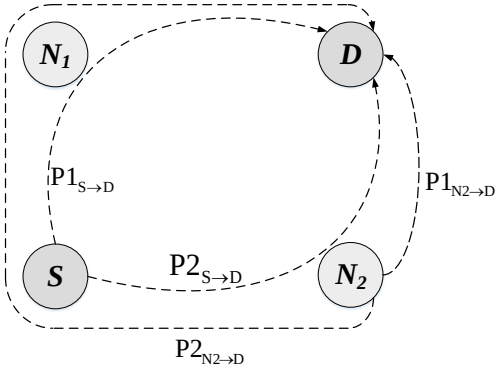


FIGURE 3.5: Possible loop creation for low priority packets.

3.4 Results and Discussion

In this section, the evaluation of the proposed modifications for the HWMP protocol is presented. As said before, the ns-3 simulator have been chosen to carry out the performance evaluation. It includes the default 802.11s model which was modified to include the proposed mechanisms. In the following sections, the scenario used for the simulations is presented, together with the simulation parameters and the obtained results.

3.4.1 Simulation details

The scenario for the evaluation of this proposal is shown in Figure 3.6. This scenario consists of a grid topology where the transmitted data traffic is bidirectional. That is, home users (smart meters and other home devices available at the HANs) transmit different applications (traffic types) to the data concentrator, such as periodic billing data (meter reading), Electric Vehicle (EV) charging information and home energy among other applications. On the other hand, the data concentrator is capable of transmitting demand response information to home users, with the aim, for example, of adjusting the energy consumption during peak hours. In the

simulations, the number of nodes in the grid is a variable parameter, the data concentrator is located in the bottom left corner, and all applications are running simultaneously in every HAN.

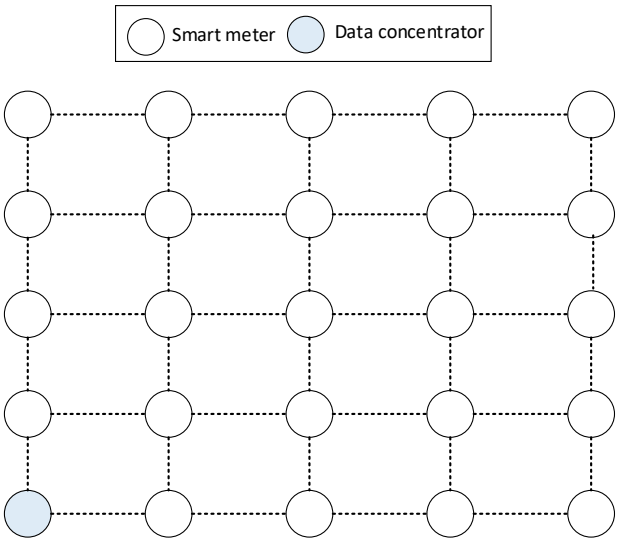


FIGURE 3.6: Scenario under consideration

Table 4.8 indicates the general parameters used in the simulations, where for each grid size (number of NAN nodes), different simulation runs have been carried out. Each run has been configured with different random seeds in order to obtain confidence intervals for the chosen performance measures: packet delivery ratio, throughput, network transit time, and routing table size. To ensure the network topology shown in Figure 3.6, 80 meter have been chosen as the grid distance between nodes. With this value, and the propagation model parameter values, each node is only able to establish connections with the neighbors located on its sides.

TABLE 3.3: General simulation parameters.

Description	**Value**
Simulator	ns-3.28
Number of nodes	from 9 to 36
distance between nodes	80 m
Simulation time	50 s
Tranport layer	User Datagram Protocol (UDP)
Random number generator	MRG32k3a

The applications that will be transmitted over the NAN network are detailed in Table 3.4, where they have been grouped into four types of traffic according to the EDCA categories and the quality of service requirements. The table shows also the distributions and average values selected for the packet size and for the interarrival time, where two main distributions have been considered. On the one hand, for periodical traffic types where applications generate packets of constant size at regular time intervals, the distributions have been considered as deterministic. On the other hand, for traffic types based on events or by variable rate applications, an exponential distribution has been selected, which can be better adjusted to the combination of this type of applications. To implement the latter distribution, the default application module of the ns-3 simulator has also been modified. In addition, it has been considered interesting to

analyze the network performance in two different load conditions. Thus, two different sets of values have been selected for the packet generation rate. Firstly, a relatively high network load (NL1) was considered, and secondly, a network load that causes an extreme congestion situation (NL2).

TABLE 3.4: Applications classification, distributions and parameters.

Applications	**Packet size**		**Packet Interval**			**EDCA**
	Length (Bytes)	**Distribution**	**Interarrival time (secs) NL1**	**NL2**	**Distribution**	
Demand Response, Outage Management.	60	Exponential	0.075	0.025	Exponential	Voice (Highest Priority: 1)
Video surveillance, Overhead Transmission Line Monitoring, Substation Automation systems (SASs).	60	Exponential	0.075	0.025	Exponential	Video (Priority 2)
Home Energy Management (HEM), Electric Vehicles (EVs) Charging	512	Deterministic	0.075	0.025	Deterministic	Background (Priority 3)
Meter Data Management	512	Deterministic	0.075	0.025	Deterministic	Best Effort (Lowest Priority: 4)

The IEEE 802.11 standard defines the methods to initiate, maintain and close the bidirectional links between mesh STAs, and also establishes by default the Hybrid Wireless Mesh Protocol and the Airtime link metric. Tables 3.5 and 3.6 present the parameters configured in the simulator for the Mesh Peering Management (MPM) and HWMP protocols, where, among others, the following variables are defined: maximum thresholds to consider invalid links, maximum number of neighbors (peer links) allowed, the reactive mode of HWMP, lifetime of the reactive routing information and the conditions to indicate a route as unreachable. As mentioned in the previous section, nodes must be allowed to establish links only with the neighbors at their sides. To reinforce this, the maximum number of peer links per node has been set to four. On the other hand, the 802.11s model implemented in the ns-3 simulator considers a link as not valid if the consecutive number of lost beacons achieves a configurable threshold (*maxBeaconLoss* in Table 3.5). A value of 20 lost beacons has been selected for this parameter. In addition, when a station is unable to transmit to its peer a number of successive data frames, the ns-3 implementation by default close their peer link. This parameter and the other variables presented in Tables 3.5 and 3.6 were configured with their default values. These selections does not affect the performance evaluation carried out, since the values are the same for both compared protocols.

TABLE 3.5: Mesh Peering Management protocol parameters.

Variable	**Description**	**Value**
maxRetries	Maximum number of retries	4
maxBeaconLoss	Maximum number of lost beacons before link will be closed	20
maxNumberOfPeerLinks	Maximum number of peer links.	4
maxPacketFailure	Maximum number of failed packets before link will be closed	5

TABLE 3.6: Hybrid Wireless Mesh Protocol (HWMP) parameters.

Variable	Description	Value
pathMode	Path selection mode	On-demand
maxQueueSize	Maximum number of packets we can store when resolving route	255
maxPREQretries	Maximum number of retries before we suppose the destination to be unreachable	5
reactivePathTimeout	Lifetime of reactive routing information	5.12 sec

Table 3.7 presents the configured values for the physical layer, detailing among others the following parameters: 802.11a as the selected physical layer, frequency channels for control and data traffic and propagation model. Except for the number of control and data channels and their frequencies (which have been defined in this proposal), well known values have been chosen for the rest of parameters in Table 3.7, which are used in most Smart Grid NAN simulations. As previously said, this selection does not affect the comparison between the protocols.

TABLE 3.7: Physical layer parameters.

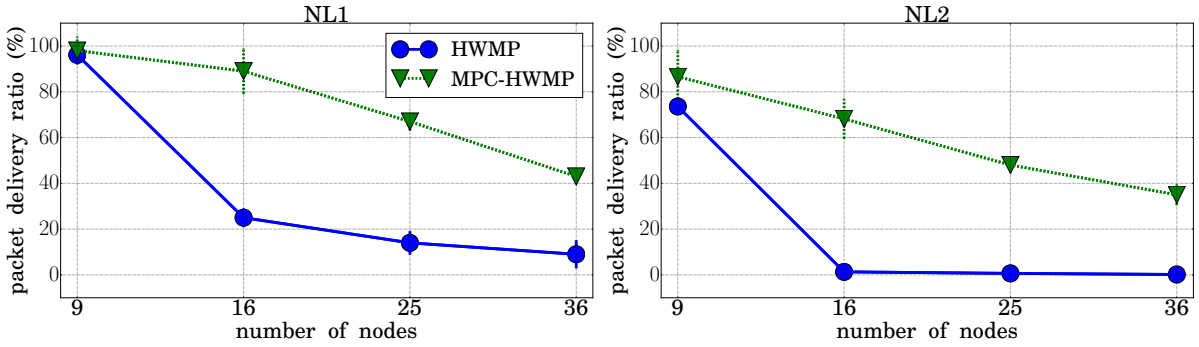
Variable	Description	Value
phyLayer	Wireless physical layer.	802.11a
controlChannelNumber	Number of control channels.	1
controlChannelFreq	Frequency of the control channel .	5180 MHz
dataChannelNumber	Number of data channels.	4
dataChannelFreq	Frequency of data channels	5200 MHz 5220 MHz 5240 MHz 5260 MHz
propagationDelay	Maximum propagation delay	3.333
propagationModel	Propagation loss model.	Log distance
	Exponent: The exponent of the Path Loss propagation model	3
	ReferenceDistance: The distance at which the reference loss is calculated (m)	1 m
	ReferenceLoss: The reference loss at reference distance (dB). (Default is Friis at 1m with 5.15 GHz)	46.667

3.4.2 Numerical results

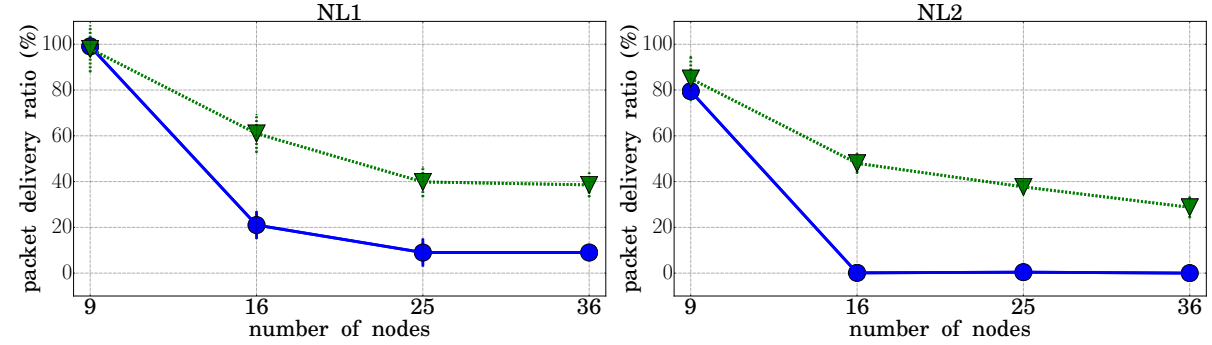
In this subsection, the obtained results are presented and evaluated. Although the ns-3 simulator provides some tools for data analysis, they are mainly designed to work with protocols that operate at the network layer. As this work is focused on a protocol (HWMP) that operates at the data link layer, a new tool has been designed. In the followings sub-sections the obtained results are evaluated in terms of packet delivery ratio, network transit time, throughput, routing tables size and control channel utilization.

Packet delivery ratio (PDR)

We compare the Hybrid Wireless Mesh Protocol (HWMP) algorithm with the proposed extension Multi-Path Multi-Channel Hybrid Wireless Mesh Protocol (MPC-HWMP) in terms of packet delivery ratio (and its 95% confidence interval) for different grid sizes (from 9 to 36 nodes). The PDR defines the relationship between the number of successfully received packets and the total number of transmitted packets. Figure 3.7 shows the results for the four traffic types considered (Figure 3.7(a),(b),(c) and (d) respectively). Besides, for every traffic type, the graph on the left side shows the results under the load conditions NL1, while the one on the right shows them for NL2. The results confirm that, as the size of the network increases, the PDR decreases for the four traffic types and for the two network load conditions considered. As can be seen, under the NL1 conditions the network starts to be very loaded for a number of nodes greater than 16, while the NL2 conditions leads to a total saturation and a PDR value equal to zero when the basic HWMP is used. However, the PDR decrement is much lower with MCP-HWMP. In addition, this figure highlights that, when using MPC-HWMP, the traffics with higher priorities, which use the best available paths, receive a better service from the network than those with lower priorities.

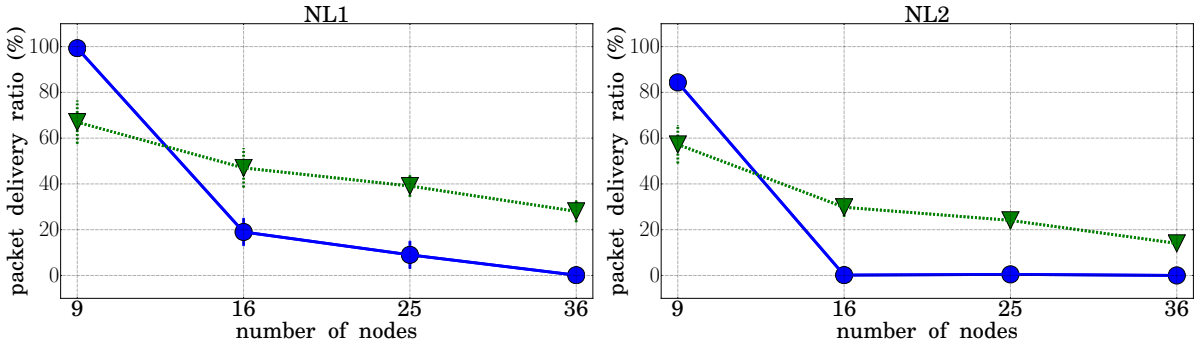


(a) Traffic type 1

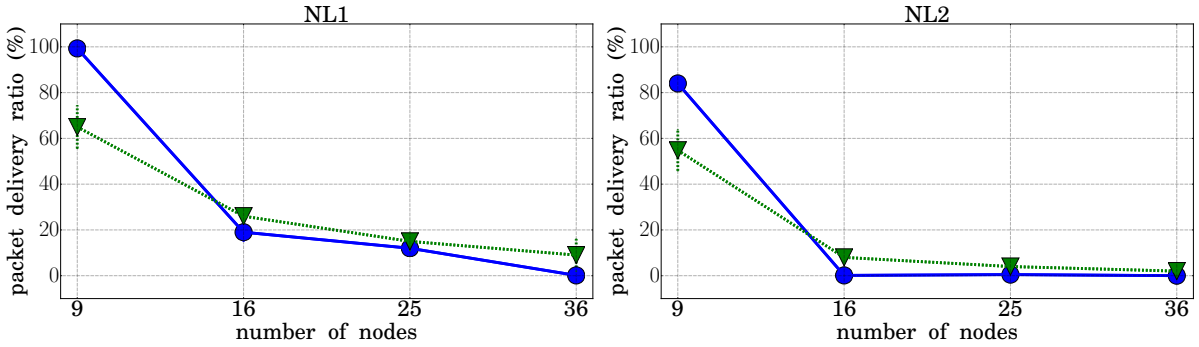


(b) Traffic type 2

FIGURE 3.7: Packet Delivery Ratio (HWMP vs MPC-HWMP) .



(c) Traffic type 3

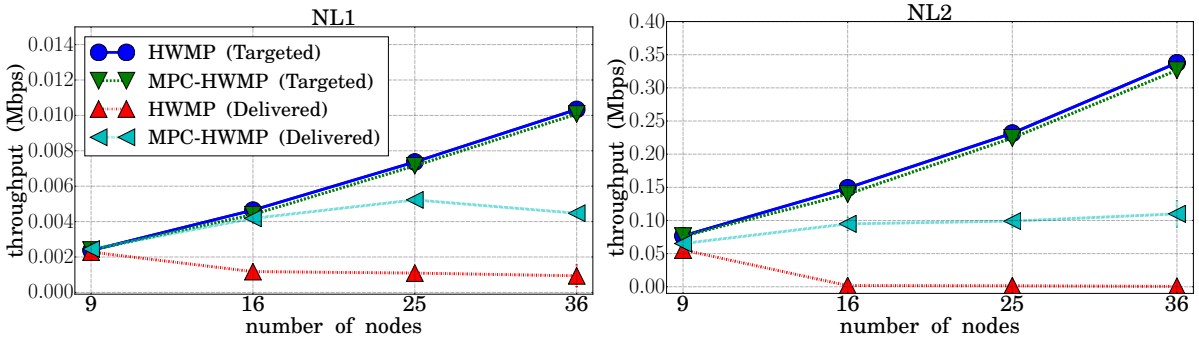


(d) Traffic type 4

FIGURE 3.7: Packet Delivery Ratio (HWMP vs MPC-HWMP) (cont.).

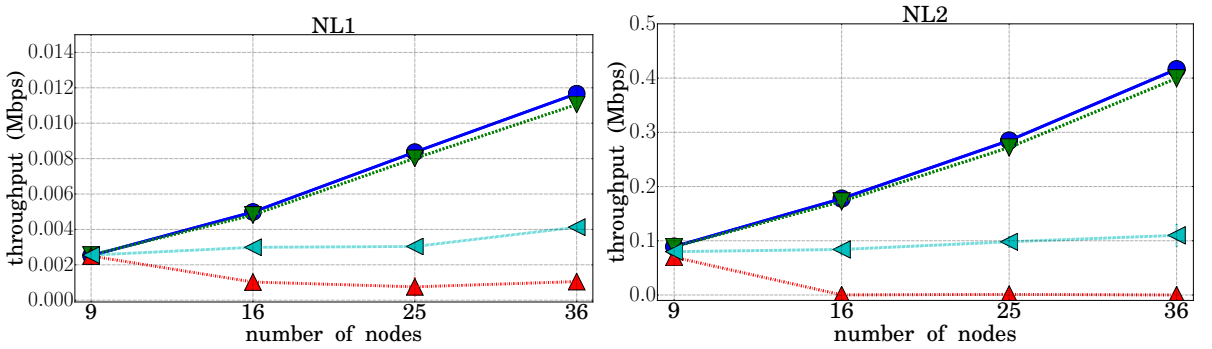
Network throughput

In this section, the results obtained for the network throughput are presented. The throughput represents the number of bits per second transmitted correctly, and it is a performance parameter that complements the PDR offered in the previous section. Figure 3.8 shows on the one hand the "targeted" throughput, which consists of the bits per second generated by all the applications. As it can be seen, this throughput is the same regardless of whether the protocol used is HWMP or MPC-HWMP, and it is higher for the NL2 load conditions. However, the throughput correctly delivered to its corresponding destination is higher when the protocol used is MPC-HWMP. In particular, it can be checked that the throughput delivered with HWMP tends to zero when the network size is equal to or greater than 16 nodes, which is consistent with the results already commented for the PDR.

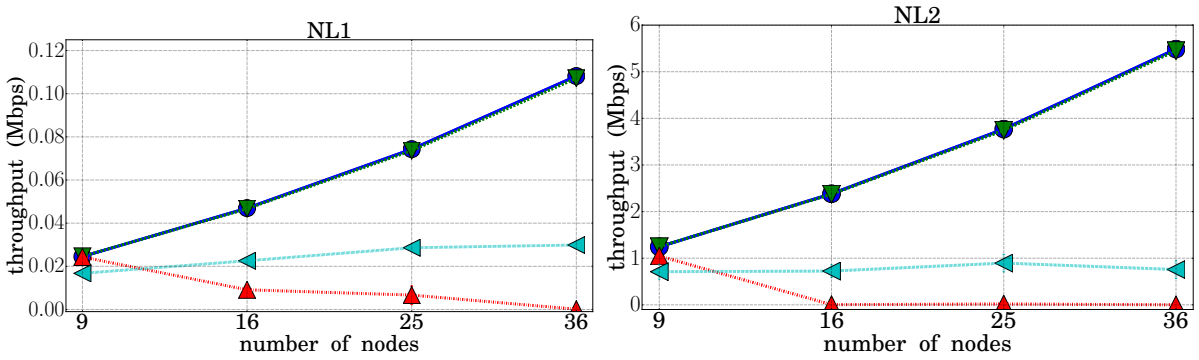


(a) Traffic type 1

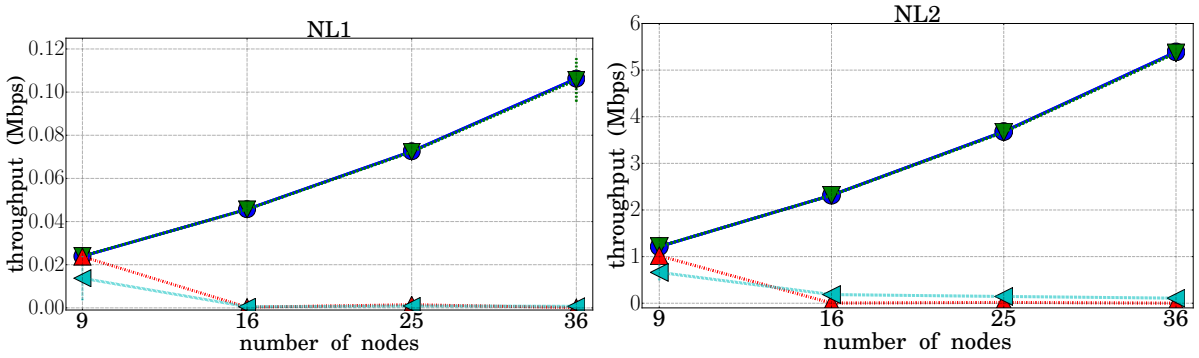
FIGURE 3.8: Throughput (HWMP vs MPC-HWMP)



(b) Traffic type 2



(c) Traffic type 3



(d) Traffic type 4

FIGURE 3.8: Throughput (HWMP vs MPC-HWMP) (cont.)

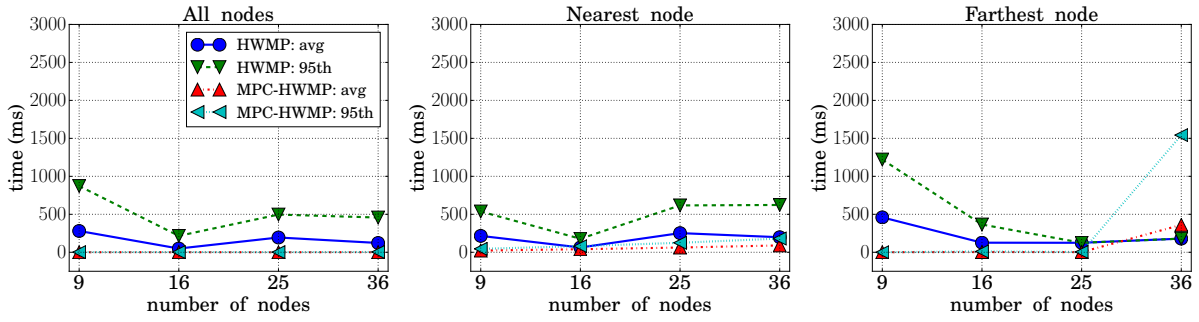
Network transit time

The network transit time is the time that packets need to go from their source to their destination through the intermediate nodes. For this parameter, instead of offering just the average value, which could hide relevant variations in the service offered to different packets, it has been considered of importance to offer also the percentile value (specifically, the 95th percentile was chosen). On the other hand, to show the existing differences depending on the specific location of each node with respect to the data concentrator, the results are provided separately for the nodes that are in the best and worst situation, that is, for the nearest node and for the farthest node to the data concentrator.

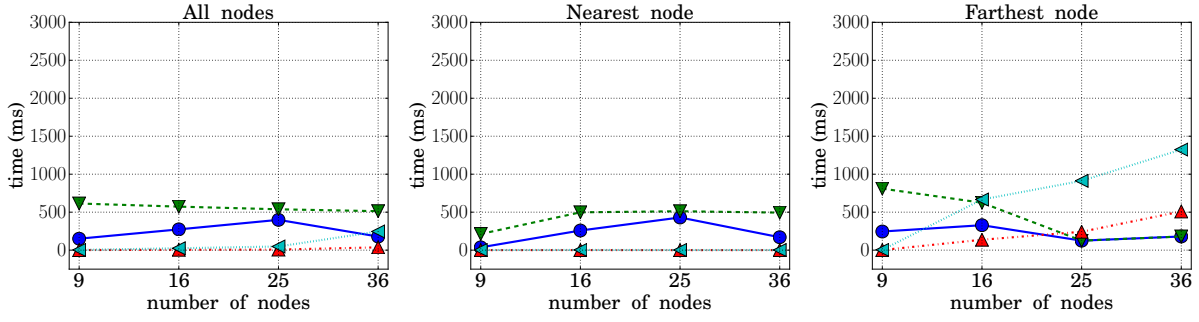
Figure 3.9 compares the average and the 95th percentile values for the HWMP and MPC-HWMP cases, when the network is working under load conditions NL2. For each traffic type, the graph on the left side shows the averaged values considering all the network nodes, the graph on the

center shows the values considering only the nearest node to the data concentrator, and the graph on the right side takes into account only the values obtained for the farthest node. The general tendency of these time values should be to grow as the network size is increased, and it can be observed that network transit times are smaller when the proposed MPC-HWMP is used. However, paying more attention to the particular cases, some details have to be discussed. First, it can be observed that in an extreme congestion situation and with the basic HWMP protocol, the transit time values shown for the farthest node become smaller instead of greater when the network size is incremented (this also occurs, although to a lesser extent, for other nodes and with NL1 load conditions). This is due to the fact that the vast majority of packets are being lost (remember that the PDR is practically zero), so that only the values of the few packets that arrive correctly are taken into account. These packets have found the network in a instantaneous (and transitory) low load situation, and therefore its transit time has been small. However, looking at the MPC-HWMP protocol, where a significantly greater number of packets have been transmitted correctly, the value of the network transit time grows with the number of nodes as expected. Here we have an exception again, since for traffic 4 (lowest priority) the high amount of losses (see Figure 3.7) cause the same effect as for the HWMP protocol.

On the other hand, looking at the nearest node, the growth in the value of the transit time as the network size is increased, is smoother because this node in particular is affected much less by the increase of the network size. For all cases, it can be observed that the best performance in terms of transit time is obtained with MPC-HWMP.

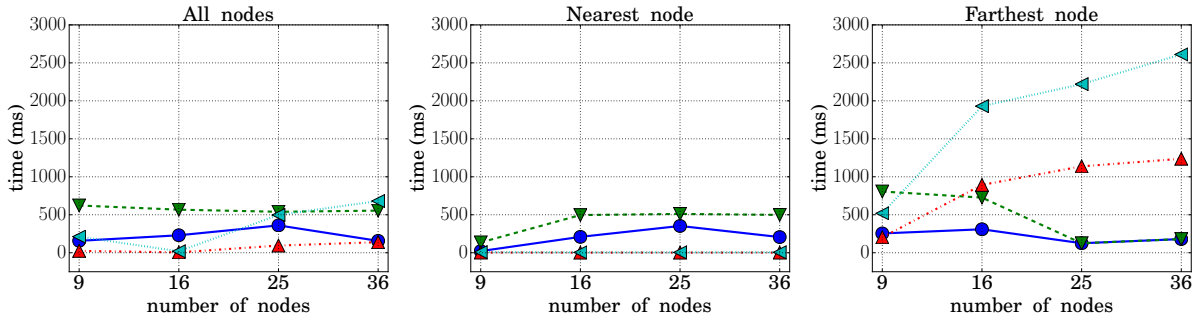


(a) Traffic type 1

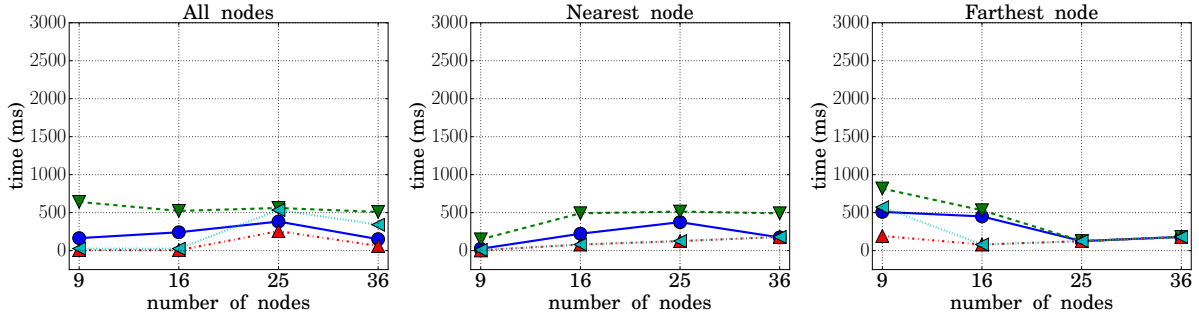


(b) Traffic type 2

FIGURE 3.9: Network transit time (HWMP vs MPC-HWMP).



(c) Traffic type 3



(d) Traffic type 4

FIGURE 3.9: Network transit time (HWMP vs MPC-HWMP) (cont.).

Routing table size

As explained above, to obtain the advantages offered by the multi-path mechanism it has been necessary to increase the number of entries that each node must store in the routing table, as well as to add some fields to those entries. Both actions translate into an increase in the amount of memory required in each node to store its routing table. To quantify this increase, the routing table size has also been measured during the simulations. Figure 3.10 shows the results, both for the basic HWMP and for the modified protocol. For every network size, the minimum and maximum values are depicted (that is, the nodes with the smallest and largest table). Besides, the boxes represent the 25th, 50th and 75th percentiles values.

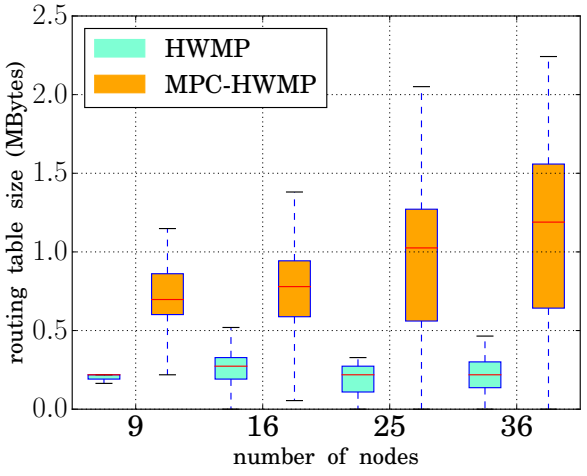


FIGURE 3.10: Routing table size (HWMP vs MPC-HWMP)

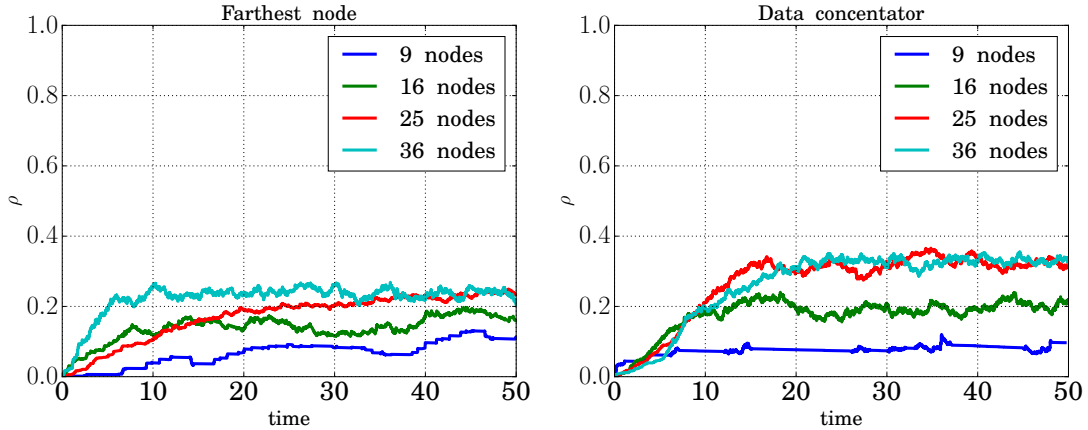
When the basic HWMP is used the routing table size is independent of the network size, since the nodes must only store the route to the data concentrator. However, with the multi-path modifications, the different available paths are stored to be assigned to the different traffic types. Thus, when the network size is increased, more paths become available and so the amount of memory needed for the routing table also grows. As shown in the figure, in the worst case taken into account in the simulations (36 nodes) the amount of memory needed (for the node with the largest table) is around 2.3 megabytes.

This fact could represent a scalability problem if the number of nodes in the network could grow indefinitely, but this is not the case with Smart Grid NANs, where one node represents one home. In any case, the amount of memory needed could reach the order of tens or hundreds of Mbytes, which with current memory technologies does not represent any problem.

Control channel utilization factor

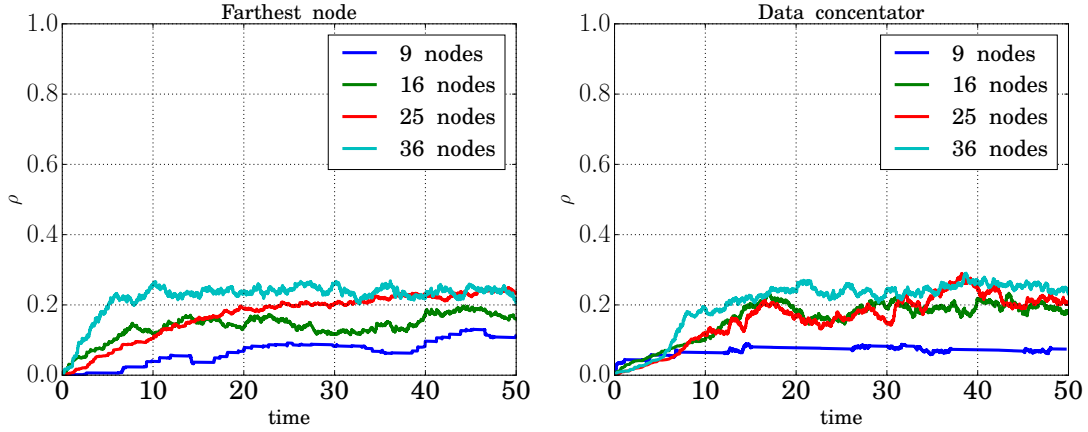
The possibility of having multiple paths implies an increase in the number of PREQ and PREP packets that must be transmitted. To avoid the saturation of the data channels, in this proposal, an exclusive channel has been assigned for control packets. In order to evaluate the impact caused by these new transmissions, the measure of the control channel utilization factor (ρ) has also been carried out. This parameter can be measured by each mesh station, being the possible values 1 (busy) or 0 (idle). These values are then smoothed using an exponentially weighted moving average (EWMA) in order to obtain an estimation of the average value and avoid abrupt oscillations.

The results are shown in the Figure 3.11, where the channel utilization factor for two cases (farthest node and the data concentrator), for two different values of the *reactivePathTimeout* parameter (lifetime of reactive routing information values), and for different network sizes (from 9 to 36 nodes), is presented. It can be observed how, as the network size is increased, more control packets are transmitted and so the ρ value also increases, for both *reactivePathTimeout* values. Besides, it is also shown that the ρ value does not achieve high values (it is always below 0.4) and therefore the control channel is not congested.



(a) Lifetime of reactive routing information: 2.05 s

FIGURE 3.11: Control channel utilization factor.



(b) Lifetime of reactive routing information: 5.12 s

FIGURE 3.11: Control channel utilization factor (cont.).

3.5 Conclusions and Future Work

In this chapter, the implementation of Smart Grid NANs with multi-hop wireless mesh networks has been considered. Specifically, a modification of the HWMP protocol has been proposed and evaluated, based on the maintenance of multiple paths between each pair of network nodes. In addition, independent frequency channels have been defined for each of the paths, as well as a special channel for control messages. With the joint application of both techniques, a more efficient utilization of the available network resources has been achieved.

To evaluate the obtained benefits, the ns-3 network simulator has been used, on which all the proposed modifications have been implemented. The results of the simulations allowed to verify the improvements in the network performance in terms of packet delivery ratio, throughput and network transit time. On the other hand, since the application of multipath techniques supposes an increase in the size of the routing tables stored in the nodes, the necessary amount of memory to store them has also been measured. The results allow to affirm that no memory problems will arise in the nodes.

As future lines of work, research to improve NANs performance will continue by adding network security and data privacy. In this field, new extensions for the HWMP protocol based on multiparametric optimization techniques will be proposed, in order to take advantage of the improvements obtained with the multipath proposal and at the same time guaranteeing a better network service by prioritizing the paths through the nodes with the better reputation.

In the following chapters, different congestion control mechanism are proposed, implemented and evaluated to improve the network performance for those Smart Grid NANs.

Chapter 4

Emergency and Fairness Aware Mechanisms for Congestion Control

Several applications are transmitted over the NAN, and some of them are strictly important for the Smart Grid operation. Besides, if those traffics are not regulated, it can lead to network congestion problems. Typical effects of network congestion are higher queueing delays, increasing of the network packet transit time, packet losses and connection losses. Theses effects can reduce drastically the network performance of NAN, and specially, for those applications with higher QoS requirements. In this chapter, two congestion control mechanisms are proposed, implemented and evaluated when the selected technologies are the wireless mesh network and the wireless ad hoc network. Besides, multi-channel allocation schemes, emergency system module, traffic differentiation and a fair distribution of network resources are covered throughout the chapter.

4.1 Introduction

Congestion control involves a set of techniques to detect and correct the problems when the whole targeted traffic can not be transmitted. That is, a link transmits more data that it can handle. In this context, network congestion problems can appear in a NAN scenario, and if those problems are not mitigated, it can reduce the network performance. Basically, the geographical position of the nodes, network size, and unregulated source rates can saturated partially or totally the NAN. Remember that most Smart Grid applications have strong security and reliability requirements, and thus, congestion control mechanisms are mandatory in order to provide QoS provision. In this chapter, two different techniques that allow improving the performance offered by the NANs in emergency and high network congestion situations are proposed, implemented and evaluated.

Firstly, a congestion control mechanism which works together with an emergency system is proposed, implemented, and evaluated. Besides, the proposed solution also applies a multi-channel allocation scheme and traffic differentiation. The solution is based on congestion control functions that based on the current network load and emergency state provides more or less transmission probability to priority traffics. For each emergency state, a set of congestion control functions is proposed. This proposal has been evaluated in the context of a wireless mesh networks made up by a set of smart meter devices, where various smart grids applications

are sending their data traffics. Basically, two scenarios are used to evaluate the solution. In the first scenario the proposed mechanism is assessed for a congestion control scenario. While, the second one evaluates both a congestion and an emergency scenario.

Secondly, a fair and distributed congestion control mechanism is considered. The solution is agnostic to the network and mac layers. In order to show the versatility of the proposed mechanism, the wireless ad hoc network, the AODV routing protocol and one of the latest physical standards (802.11ac) are now used. Similar to the previous solutions, traffic differentiation is also provided for those critical applications with higher QoS needs. Besides, the distributed solution also provides a fair distribution of the network resources. The distributed solution is evaluated for three scenarios. The first simple scenario explains the proposed solution, and the other scenarios evaluates the solution in a Smart Grid NAN environment.

The presented techniques are executed individually in every node, and do not significantly load the node's CPU, which is advisable if we take into account that in many cases these devices are built on a large scale and at low cost, and so they have limited resources. Both proposed solutions are evaluated in the ns-3 simulator. Applying our proposed congestion control mechanisms lead to performance improvements in terms of packet delivery ratio, network throughput fairness between different traffic sources, packet network transit time and QoS provision.

The rest of the chapter is organized as follows. In Section 4.2 we report and analyze the related work. Section 4.3 presents the emergency aware congestion control mechanism and Section 4.4 presents the fair and distributed congestion control mechanism. Finally, the conclusions and future works are summarized in section 4.5.

4.2 Related works

Several researchers have focused their work on the proposal of new mechanisms, or on the modification of existing ones, with the aim of improving the performance offered by wireless multihop networks in smart grid neighborhood area networks. In WMN, there are two types of congestion: intra and inter-mesh congestion. Although there are multiple algorithms to solve intra-mesh congestion, in the context of NAN networks there are no works that focus on congestion control, since they are, in general, aimed to WMN. For instance, a congestion control mechanism for WMN is presented in [84]. Authors have considered that the increase in the waiting time to access the wireless channel also increase the packet delay and then, the resulting queue length leads to congestion. They propose a modification to the default HWMP in order to provide congestion avoidance. They consider that each node monitors its queue length for each flow and they notify to their neighbors when it reaches a specific level through Congestion Control Notification Frames (CCNF). This action allows the neighboring node to calculate an alternative path depending on the queue length and also excluding the congested link.

The implementation of multi-channel in WMNs can be done in two ways: multiple radios on the physical layer (PHY) or using the channel switching capability of the device. The proposals and works in [85–89] describe techniques for channel assignment, multiple beam-antennas and multiple-radio routing metrics. However, this concept, together with congestion control techniques, has not been fully investigated in NAN networks when the WMN is implemented as the technology for data communications. For instance, in a previous work [53] a basic congestion control mechanism for WMN is proposed, which works together with a multi-channel allocation scheme. In this work, two traffic types have been defined (priority and non-priority), and the

congestion control mechanism discards non-priority applications when a given channel utilization factor value is exceeded. Besides, in order to reduce the network congestion, two different channels have been used for transmitting the different NAN applications.

In this chapter, two congestion control mechanisms are proposed, implemented and evaluated, and they are explained in the following sections.

4.3 An Emergency Aware Congestion Control mechanism

In this section, a congestion control mechanism which takes into account different traffic priorities, emergency states, and a multi-channel allocation scheme is presented. Besides, three different network states have been defined (normal, medium or high), and they can be triggered manual or automatically. In this context, the nodes measure periodically different network parameters, and in case of network anomalies, they notify the rest of the nodes by using special control frames. These messages modify the congestion control mechanism in order to give more or less transmission probability to priority traffic types.

In order to implement this proposal, the wireless mesh network have used as the selected technology for NAN where the default ns-3 module has been modified.

4.3.1 Proposed solution

As previously said, the implementation of the HWMP protocol integrates these main processes: Route Management, Data Queue and Route Assignment [57]. In order to improve the performance of the basic protocol, in this subsection four specifics mechanisms, which work collaboratively, are proposed, implemented and evaluated: traffic differentiation, multi channel allocation, congestion control and emergency system. Figure 4.1 shows the structure of the modified HWMP, with the new modules and their relation with the basic ones, and Tables 4.1 and 4.2 present the definition of the variables and functions for the proposed emergency aware congestion control mechanism (EA-HWMP).

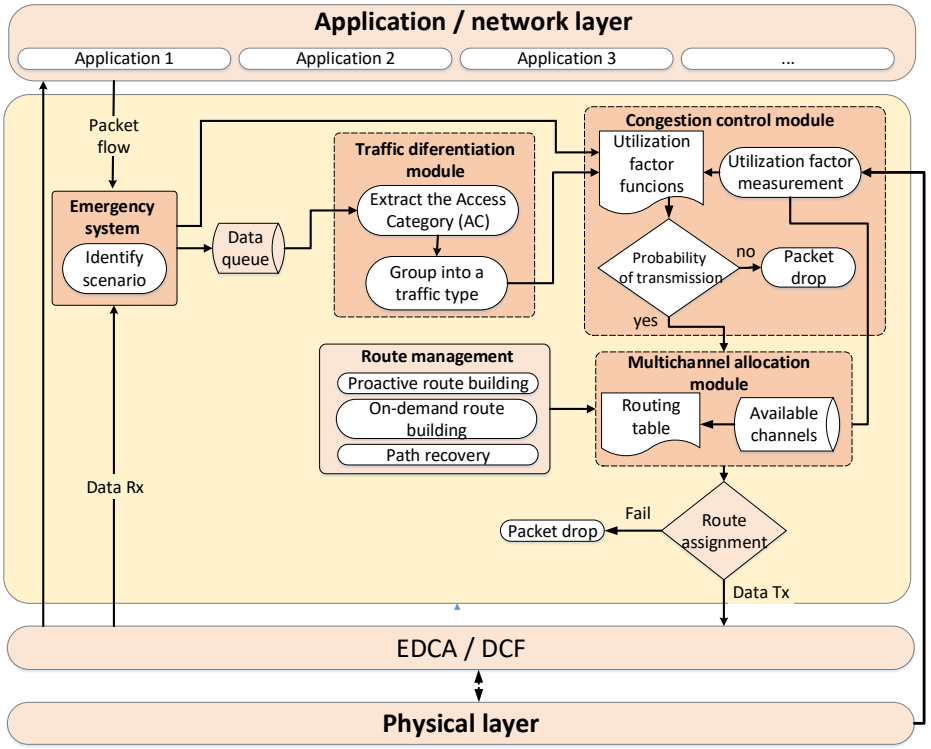


FIGURE 4.1: Structure of the emergency aware congestion control mechanism (EA-HWMP).

TABLE 4.1: Definition of the variables for the emergency aware congestion control mechanism (EA-HWMP).

Parameter	Description
N_c	Number of available data channels.
ρ	Channel utilization factor for each data channel.
ρ'	The previous value of ρ.
s_ρ	Current sample of the physical channel state.
ω	The value of $\omega(\omega \in [0,1])$ defines the importance of s_ρ with respect to the past values.
f_r	Data frame (NAN application).
ρ_T	The sum of ρ of all available data channels.
t_T	NAN applications are classified into four traffic types.
ccf	Congestion control functions.
P_T	Transmission probability based on the result of ccf.
E_S	Emergency situation (normal, medium and high).
s	Source mac address.
d	Destination mac address.
AC	EDCA access categories (voice (VO), video (VI), best effort (BE) and background (BK)).
TTL	Time-to-live.
ρ_{th}	The maximum threshold of ρ allowed per channel.
ch	Data physical channel.

TABLE 4.2: Definition of the functions for the emergency aware congestion control mechanism (EA-HWMP).

Function	**Description**
isStateBusy	Measures the physical channel state (busy: 1, iddle: 0).
mapTraffic	NAN applications are mapped to the four EDCA access categories.
shouldTransmit	Depending on the ρ_T, ccf and E_S values, the frame will be pass to the multichannel allocation module.
congestionControl	Congestion control mechanism (Algorithm 5).
channelAllocation	Data frames are assigned to a specified channel based on the value of ρ_{th} (Algorithm 7).

Different NAN applications are (re)transmitted through the NAN network, as well as routing and emergency messages. Similar to the previous proposal, the data messages (NAN applications) are grouped into four traffic types according to their priority, where the traffic type 1 corresponds to the NAN applications with the higher QoS requirements, while the traffic type 4 represents the applications with less QoS needs. Furthermore, different emergency messages are used to characterize different anomalous situations present in the smart grid (weather conditions, malicious agents, terrorist attacks, hardware or software failure, etc.). Keep in mind that the emergency messages modify the operation of the congestion control module. This module basically gives a higher probability of transmission to priority traffic types in situations of network congestion, and this probability will increase or decrease according to the current emergency state (normal, medium and high) of the smart grid. Finally, a multichannel allocation module is also implemented, which assigns a dedicated channel for the control traffic and the rest of channels are assigned to transmit the NAN applications. These modules will be explained in greater detail in the following subsections.

4.3.1.1 Route management

This module is responsible for computing and updating the paths through a proactive or an on-demand path-building mechanisms. It must be kept in mind that different channels have been used for control and data traffic, where just one channel is assigned to the control traffic (routing and emergency messages), and the rest of the available channels are assigned to transmit the data frames.

4.3.1.2 Traffic differentiation module

As it was mentioned in the previous chapter, there are different applications (data traffic flows) that are transmitted through the network are classified into four traffic types. Data packets arriving from the network / application layer are labeled with a specific access category. The objective is to map the traffic types according to the Enhanced Distributed Channel Access (EDCA) categories.

In the same way, the traffic type 1, which has the highest QoS requirements, is mapped to the voice (VO) access category, while the traffic type 4, which has the lowest QoS needs, is assigned to the background (BK) access category. The proposed scheme does not modify the medium access mechanism, since the access categories are used to identify the NAN applications and group them into different traffic types according to their priorities.

4.3.1.3 Congestion control and emergency system modules

The network congestion mechanism (Algorithm 5) is based on the value of the channel utilization factor (ρ). The ρ measurement is performed before the multi channel allocation takes place. This is because the selection of the channel for a frame to be forwarded will depend on the access category (traffic type (t_T)) and the result of the network control congestion mechanism at that precise moment.

Algorithm 5: Congestion control algorithm.

Input: data frame (f_r) (NAN application)
Output: According to the channels utilization factor functions, the frame must be transmitted or not.

for $i \leftarrow 1$ **to** N_C **do**
  $s_{\rho_i} \leftarrow \text{isStateBusy}(ch_i)$
  $\rho_i \leftarrow \omega \cdot s_{\rho_i} + (1-\omega) \cdot \rho_i'$
$\rho_T \leftarrow \sum_{i=1}^{N_C} \rho_i$
$t_T \leftarrow \text{mapTraffic}(f_r)$
$P_T \leftarrow \text{ccf}(t_T, \rho_T, E_S)$
if shouldTransmit(P_T) **then**
  return **true**
else
  return **false**

The proposed congestion control mechanism is based on the definition of a congestion control function (ccf) that assigns a transmission probability (P_T) to each value of the utilization factor. In this way, each time a node receives a packet to be (re)transmitted, this transmission will be effectively made with probability P_T, or the packet will be discarded with probability $(1 - P_T)$. On the other hand, since it is desired to differentiate the quality of service offered to the different types of traffic, it is proposed to use a different congestion control function for each of them, as shown in Figure 4.2. In addition, since the proposed mechanism is multi-channel (with a maximum of N_C available channels), the sum of the utilization factors of all available channels will be taken into account ($\rho_T = \rho_1 + \cdots + \rho_{N_C}$). Thus, in situations of network congestion (high values of ρ_T), traffic types with lower needs ($t_T = 3$ or $t_T = 4$) will be more likely discarded. It must be kept in mind that the applications assigned to traffic types with lower priority must be those that are less sensitive to possible packet losses (for example, meter reading applications that repeat and periodically retransmit the power consumption measurements).

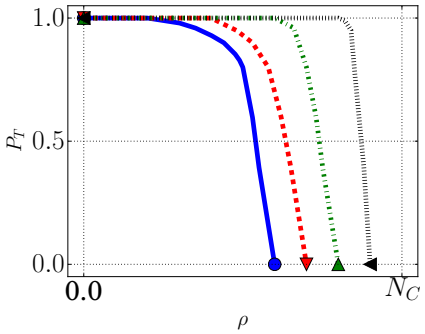
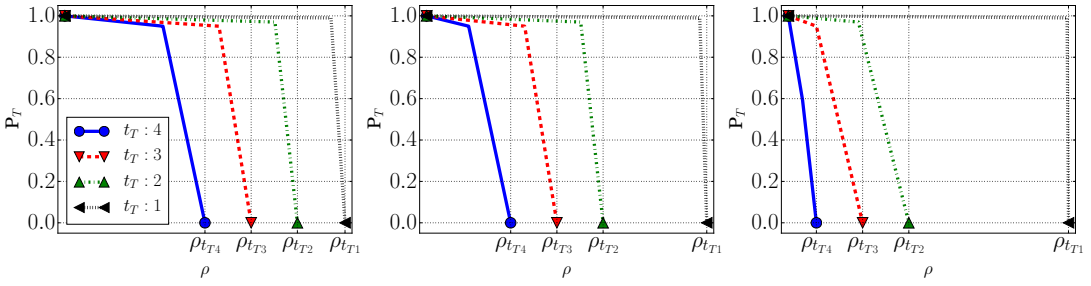


FIGURE 4.2: Congestion control function.

The congestion control mechanism works collaboratively with the emergency system module. The objective of the emergency system is to increase even more the probability of transmission of higher priority traffic in medium or high emergency situations. For this, NAN applications with lower priority are discarded at a higher rate in the time intervals in which the emergency occurs, and those with higher priorities are provided with an even better QoS. The congestion control functions, and the channel utilization factor thresholds, for each traffic type and for each emergency situation (E_S), are shown in Figure 4.3 and Table 4.3 respectively. On the other hand, the dissemination of emergency situations is done through broadcast messages, since all the nodes must know the situation. Each message contains the address of the originator node and the emergency situation, such as normal, medium or high. Finally, these messages are propagated in the network through a dedicated physical channel, to avoid high delays in congestion situations.



(a) normal. (b) medium. (c) high.

FIGURE 4.3: Congestion control functions for different emergency situations.

TABLE 4.3: Channel utilization factor thresholds for each traffic type and for each emergency situation.

	Emergency situation		
	Normal	**Medium**	**High**
$\rho_{t_{T1}}$	2	2	2
$\rho_{t_{T2}}$	1.66	1.26	0.86
$\rho_{t_{T3}}$	1.33	0.93	0.53
$\rho_{t_{T4}}$	1	0.6	0.2

After the data frame is classified according to the traffic type (traffic differentiation module), a random number is generated between 0 and 1, and if this value is below the transmission probability (P_T), the data packets are passed to multi-channel allocation module to be transmitted. Otherwise, the data packets are discarded.

4.3.1.4 Route assignment and multi channel allocation

The general route assignment tasks performed by the network nodes every time they have to (re)transmit a packet are detailed in Algorithm 6. First, the node extracts the following parameters from the packet header: source node, destination node, access category, and time to live. Then, the algorithm looks up if there is an available route. In the affirmative case, the next hop node is obtained by means of the route selection algorithm by taking into account the destination node. After, the congestion control mechanism computes if the frame should be

transmitted or not based on the emergency situation (E_S) and the network congestion (ρ_T and ccf). Finally, the transmission is assigned to a specific channel. In the case that there is no route, the packet is queued, and the path discovery mechanism is activated. This mechanism tries to obtain the route a fixed number of times, and if that threshold is exceeded, the route to that destination is considered invalid and the packet is eliminated.

Algorithm 6: Route assignment.

Input: NAN application (f_r)
Output: Forward the data to the next hop depending on the multiple channels available at the node.

```
receive_data(f_r)
[s, d, AC, TTL] ← read(f_r)
[next_hop, isThereRoute] ← RouteSelection(d)
if isThereRoute then
    if congestionControl(f_r) then
        ch_d ← channelAllocation(f_r)
        send(f_r, next_hop, ch_d)
    else
        drop(f_r)
    return
else
    if shouldInitiatePathDiscovery(d) then
        lastSQN ← getLastSQN(d)
        preq ← createPreq(lastSQN, d)
        sendPreq(preq)
    queuedPacket(packet)
```

The multi channel allocation module (Algorithm 7) is composed of two parts: the routing table created by the path-building mechanisms (proactive or on-demand) and the number of physical channels available in the current mesh STA. The purpose of this module is to select the next hop address to forward a data frame, and depending on the result of the network congestion module and the value of ρ of each channel, the transmission will be assigned to an specified physical channel.

Algorithm 7: Multichannel allocation algorithm.

Input: Number of available data channels.
Output: Assigned channel for transmission

```
channel ← 1
for ch ← 1 to N_C do
    if ρ_i < ρ_th then
        return ch
    else if i == n then
        return noChannel
```

If the output of the congestion control module indicates that the frame should be passed to the MAC sub-layer to be transmitted, the next process is to select the physical channel for the transmission. Our proposal for multichannel allocation is based on the use of the least possible number of channels at any given time. For this, the utilization factor of each channel (ρ_i) is taken into account. By default, the data packet is assigned to the data channel number 1, if and only if the value of ρ for this channel is below a certain threshold ($\rho_1 < \rho_{th}$). Otherwise, the transmission is assigned to channel number 2, and if this channel is also busy ($\rho_2 > \rho_{ch}$), the

transmission is assigned to the next channel and so on up to the maximum number of available channels (N_C). In addition, if the last available channel is also busy, the data packet will be discarded. Finally, to avoid abrupt fluctuations due to channel changes, a hysteresis cycle is taken into account around the central value of ρ_{th}.

4.3.2 Results and Discussion

4.3.2.1 Simulation details

For the numerical results, the Smart Grid NAN scenario (see again Figure 3.6) has been used again to evaluated the proposal, where various applications (traffic types) are transmitted upstream from the smart meters (SM) towards the data concentrator, and downstream from the concentrator towards the smart meters. Table 4.4 presents the selected values for the NAN applications transmitted over the NAN. Finally, different network sizes have been used to asses the current proposal. Table 4.4 shows the application parameters for each traffic type. Note that, to evaluate the network performance under stress situations, the packet generation rates have been selected relatively high for all types of traffic.

TABLE 4.4: Traffic types for different NAN applications.

Applications	**Packet size**		**Packet interval**		**EDCA**
	Length (Bytes)	**Distribution**	**Interarrival time (secs)**	**Distribution**	
Demand Response, Outage Management	200	Exponential	0.1	Exponential	Voice (Highest priority: 1)
Video surveillance, Overhead Transmission Line Monitoring, Substation Automation systems (SASs)	200	Exponential	0.1	Exponential	Video (Priority: 2)
Home Energy Managament (HEM), Electric Vehicles (EVs) Charging	400	Deterministic	0.1	Deterministic	Background (Priority: 3)
Meter Data Management	400	Deterministic	0.1	Deterministic	Best effort (Lowest priority: 4)

To evaluate the performance of the proposed mechanisms, the ns-3 802.11s basic model was again modified to build the traffic differentiation, congestion control, multichannel allocation and emergency mechanisms. The simulation setup illustrates the bi-directional flow of information between smart meters and the data concentrator using the WMN as the communication medium.

Table 4.5 presents the different values used for the application, MAC and physical layers where well-known values were chosen.

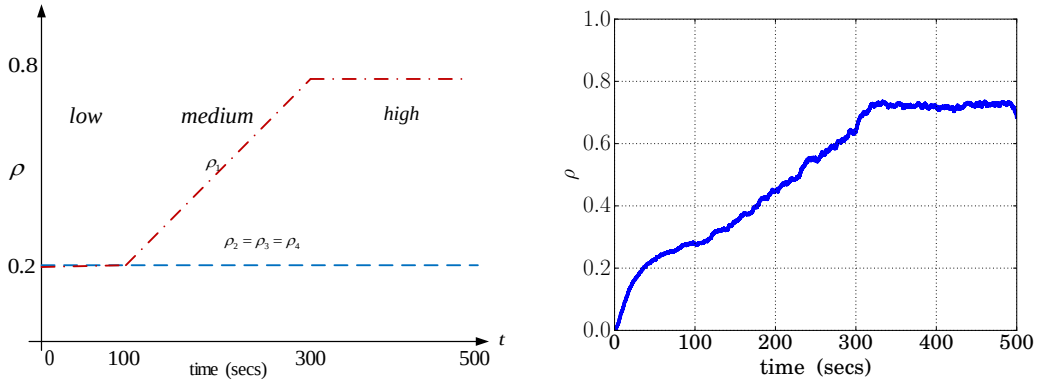
Table 4.5: General simulation parameters.

	Variable	Description	Value
Application parameters	simulator	Network simulator.	ns-3.28
	numNodes	Number of nodes.	from 9 to 36
	distanceNodes	Distance between nodes.	80 m
	simTime	Simulation time.	500 s
	tranportLayer	Tranport layer.	UDP
	randomGenerator	Random number generator	MRG32k3a
Hybrid Wireless Mesh Protocol (HWMP) parameters	pathMode	Path selection mode.	On-demand and Proactive
	maxQueueSize	Maximum number of packets we can store when resolving the route.	255
	maxPREQretries	Maximum number of retries before we suppose the destination to be unreachable.	5
	reactivePathTimeout	Lifetime of reactive routing information.	5.12 s (Case 1) 0.512 s (Case 2)
	proactivePathRootTimeout	Lifetime of proactive routing information.	5.12 s
	proactiveRootInterval	Interval between two successive proactive PREQs	1.024 s
Mesh Peering Management (MPM) protocol parameters	maxRetries	Maximum number of retries.	4
	maxBeaconLoss	Maximum number of lost beacons before the link will be closed.	20
	maxPeerLinks	Maximum number of peer links.	4
	maxPacketFailure	Maximum number of failed packets before the link will be closed.	5
Physical layer parameters	phyLayer	Wireless physical layer.	802.11a
	controlChannelNumber	Number of control channels.	1
	dataChannelNumber	Number of data channels.	3
	controlChannelFreq	Frequency of control channel.	5200 MHz
	dataChannelFreq	Frequency of data channels.	5220 MHz 5240 MHz 5260 Mhz
	ρ_{th}	The maximum threshold of ρ per channel	0.5
	propagationDelay	Maximum propagation delay.	3.333 s
	propagationModel	Exponent: the exponent of the path loss propagation model.	3
		ReferenceDistance: the distance at which the reference loss is calculated (m).	1 m
		ReferenceLoss: the reference loss at the reference distance (dB) (the default is Friis at 1 m with 5.15 GHz).	46.667

In the following sub-sections, two simulation scenarios (congestion and emergency) are implemented and evaluated in terms of packet delivery ratio, network transit time, throughput and channel utilization factor measurements. The first three parameters highlight the advantages obtained when the multichannel allocation and congestion control mechanisms are implemented, while the last one measures the channel effects when using multiple channels for the transmission of data and control frames.

4.3.2.2 Congestion control scenario

First, the behavior of the proposal is evaluated according to different network load situations, without taking into account the emergency mechanism. For this, the packet rate generation of traffic types 2, 3 and 4 remains constant, while the rate of traffic type 1 (the highest priority traffic) is increased along the simulation. This implies a channel utilization factor by each of the traffic types as shown in Figure 4.4a. The precise value of the total channel utilization factor measured by one of the network nodes during the simulation is shown in Figure 4.4b. As can be seen, in the "low network load" stage all the NAN applications have a low packet generation rate in order to not saturate the network. After a period of time, applications that belongs to the traffic type 1 begin to increase the number of the data flows ("medium network load"). In the last stage ("high network load"), the value of ρ_1 remains constant again. As previously said, in order to isolate and validate the traffic differentiation achieved by the congestion control mechanism, without considering the network emergency situation, the network state has been set as normal (see Figure 4.3a) for all this set of simulations.



(a) target ρ.

(b) measured ρ.

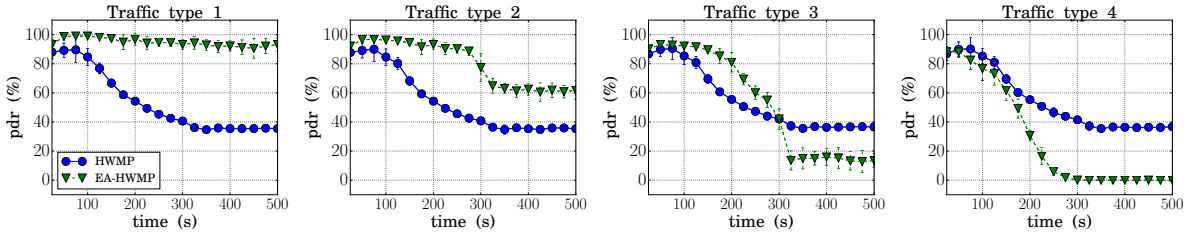
FIGURE 4.4: Traffic generation for the different NAN applications.

Packet delivery ratio and network throughput

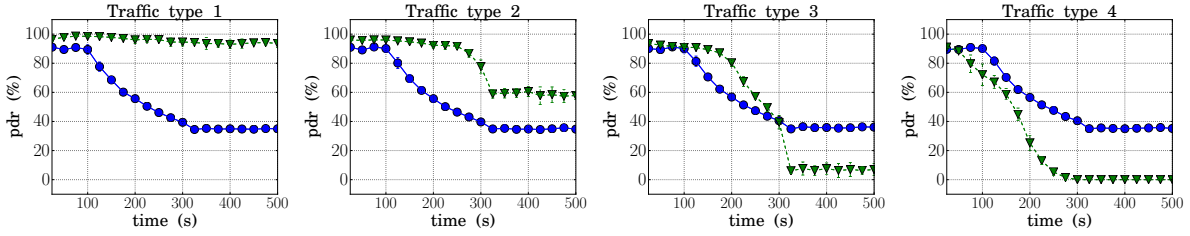
As it was mentioned in the previous chapter, the packet delivery ratio (PDR) describes the relationship between the number of successfully received packets and the total number of transmitted packets. On the other hand, the throughput represents the number of bits per second received correctly, and is a performance parameter that complements the PDR. The benefits obtained in terms of PDR and throughput are presented in two ways. First, the HWMP and our proposal (EA-HWMP) are compared taking into account the evolution of the PDR and the throughput over time. Second, the average values are compared for different network sizes.

As said, in order to check the correct behavior of the congestion control mechanism, the temporary evolution of the PDR and the throughput (and their 95% confidence interval) are shown in Figures 4.5 and 4.6 respectively, in a different plot for each traffic type. The network size chosen to show these results is 16 nodes, and three HWMP configurations have been considered. These different settings has been selected in order to show how similar are the on-demand and proactive modes for HWMP for this scenario. As it be shown later in the results, the different HWMP modes does not reproduce different results. However, a more detailed explanation will be presented at the end of this subsection.

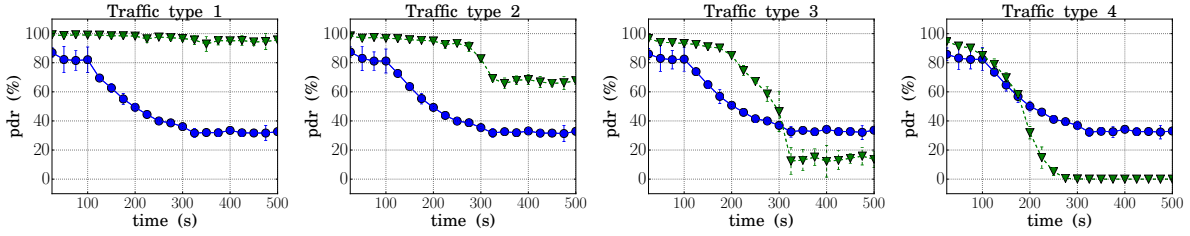
As can been observed in Figure 4.5, with the basic HWMP protocol the behavior is the same for the three HWMP configurations and for all traffic types, with a decreasing PDR as the network load grows (Figure 4.4). However, EA-HWMP prioritizes the transmission of traffic types with higher QoS needs when the network load grows. On the other hand, in Figure 4.6, the throughput (bits per second correctly received) is compared with the target throughput (bits per second transmitted) for both protocols. Obviously, the target throughput is the same regardless of whether the protocol used is HWMP or EA-HWMP. Nevertheless, the throughput is always higher when the protocol used is EA-HWMP for priority traffic types. In this case, when a frame is ready to be forwarded, the mesh STA looks up in its routing table the next hop address and depending on the congestion control functions (Figure 4.3), which are based on the sum of ρ of all available data channels and the traffic type, a transmission probability is calculated. As it was explained, the congestion control functions give more importance to traffic with greater quality of service requirements in situations of network congestion. Then, if the frame is selected for transmission, a specified channel will be assigned considering the maximum value of ρ allowed per channel. That is, if the ρ threshold (ρ_{th}) is exceeded, the next available channel will be assigned to transmit the frame. In addition, Figure 4.6, shows how the delivered throughput for traffic type 1 is increased in the "medium network load" stage, while the delivered throughput for the other NAN applications starts to decrease.



(a) On-demand mode. Reactive routing information lifetime: 5.12 s.



(b) On-demand mode. Reactive routing information lifetime: 0.512 s.

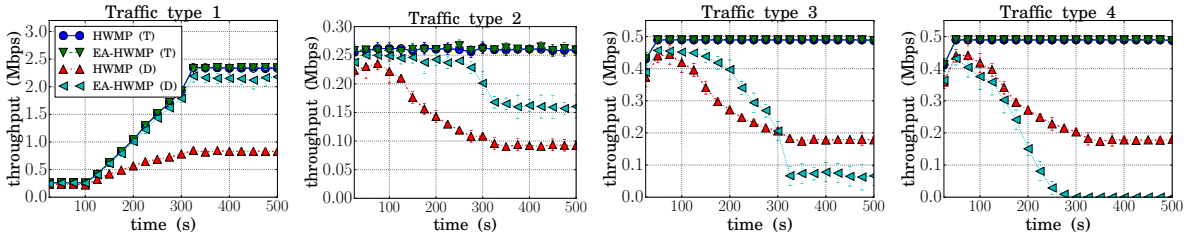


(c) Proactive mode. Proactive routing information lifetime: 5.12 s.

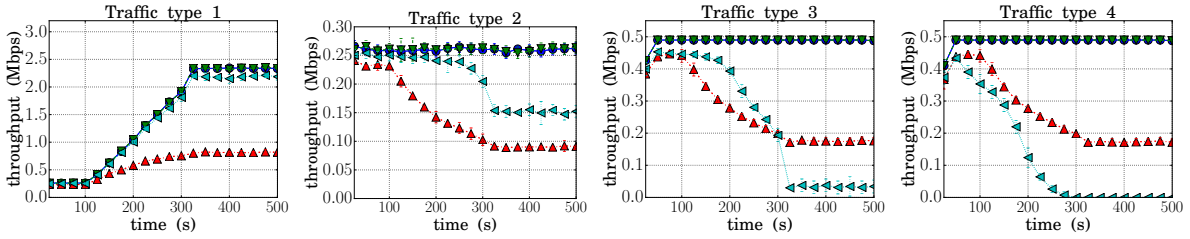
FIGURE 4.5: Packet delivery ratio (temporary evolution, grid size: 16 nodes).

Besides, Figure 4.5 and 4.6 show the same results for the on-demand and proactive modes when working in a static scenario. It must be kept in mind that these two modes differ mainly in the mechanism to disseminate the PREQ messages. In the on-demand mode, PREQ messages are transmitted when a route to a destination node expires (*reactivePathTimeout* from Table 4.8). While in the proactive mode, these messages are sent based on the time interval value

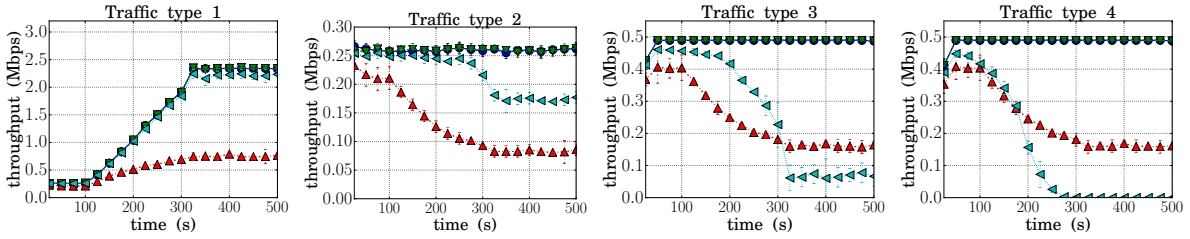
configured between two successive proactive PREQs (*proactiveRootInterval* from Table 4.8). Therefore, modifying the *reactivePathTimeout*, *proactiveRootInterval*, and *proactivePathRootTimeout* variables does not modify the paths calculated by the algorithm. In order to reduce the amount of network resources used to transmit routing messages, the following results will be shown for the on-demand mode, and when the *reactivePathTimeout* variable is set to 5.12 s.



(a) On-demand mode. Reactive routing information lifetime: 5.12 s.



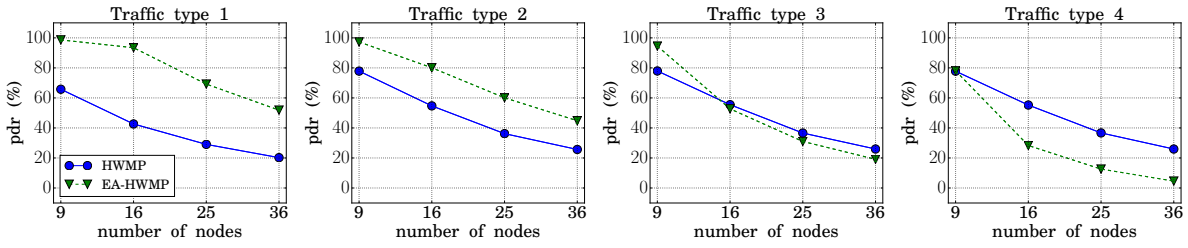
(b) On-demand mode. Reactive routing information lifetime: 0.512 s.



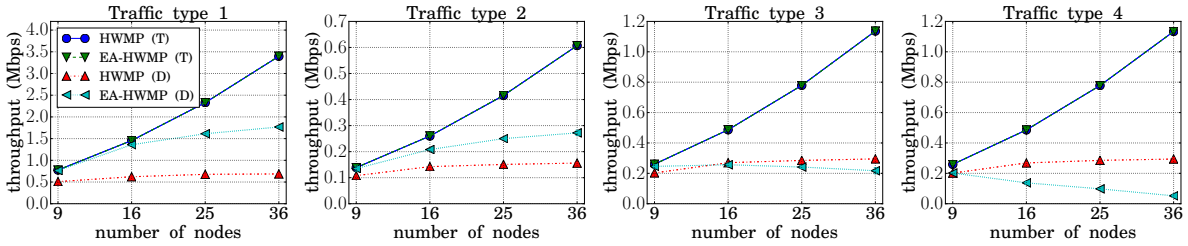
(c) Proactive mode. Proactive routing information lifetime: 5.12 s.

FIGURE 4.6: Network throughput (temporary evolution, grid size: 16 nodes).

The second set of results is shown in Figure 4.7, where the average PDR and throughput (and their 95% confidence interval) for different network sizes (from 9 to 36 nodes), and for the four traffic types under consideration, is presented. It can be observed how, as the network size is increased, it becomes more congested and so the PDR decreases for all traffics. However, this decrease is smaller with EA-HWMP for priority traffic types. Of course, the price to pay is a greater decay in the PDR for lower priority traffics, and this decrease is even greater when the number of nodes is increased in the simulations. Therefore, there is a trade-off to guarantee a targeted PDR for priority traffic at the expense of the loss of a percentage of non-priority traffic. As already mentioned, only those applications which are less sensitive to packet loss must be included in lower priority traffic types. Finally, as Figure 4.7b shows, the target throughput increases with the size of the network, since a greater number of nodes implies a greater volume of transmitted traffic (keep in mind that all the nodes transmit equally). Again, when EA-HWMP is used, for higher priority traffics the difference between the value of the target and actual throughput is lower, at the expense of a greater difference for lower priority traffics.



(a) Packet delivery ratio.

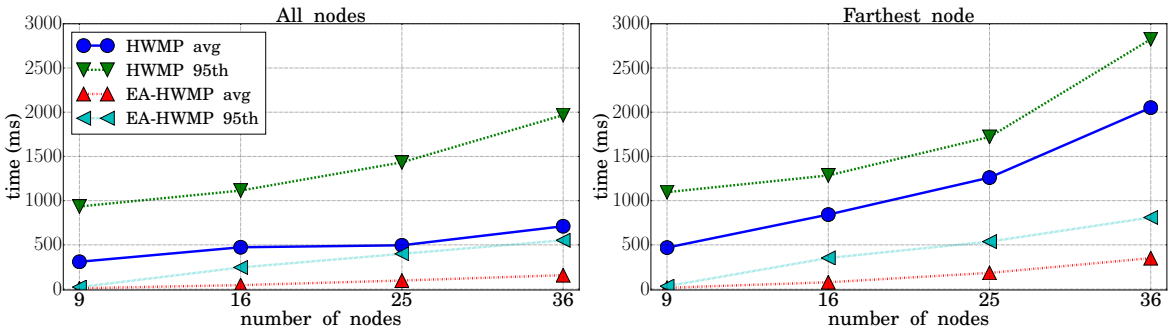


(b) Network throughput.

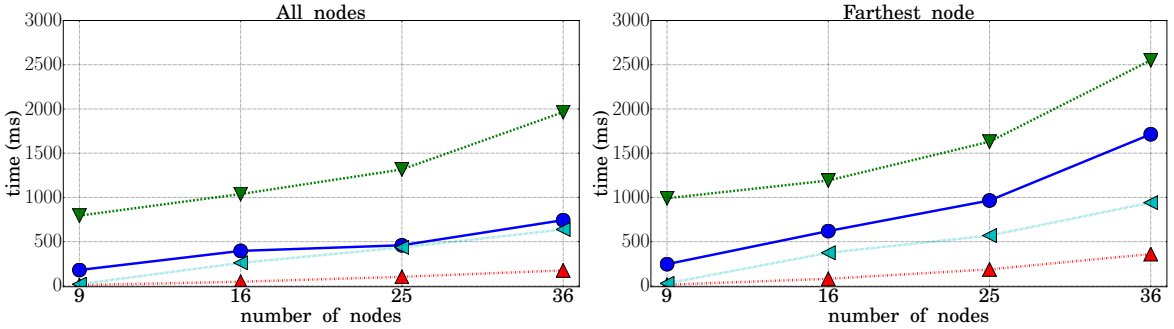
FIGURE 4.7: Packet delivery ratio and throughput for different network sizes. Reactive Mode. Reactive routing information lifetime: 5.12 s.

Network transit time

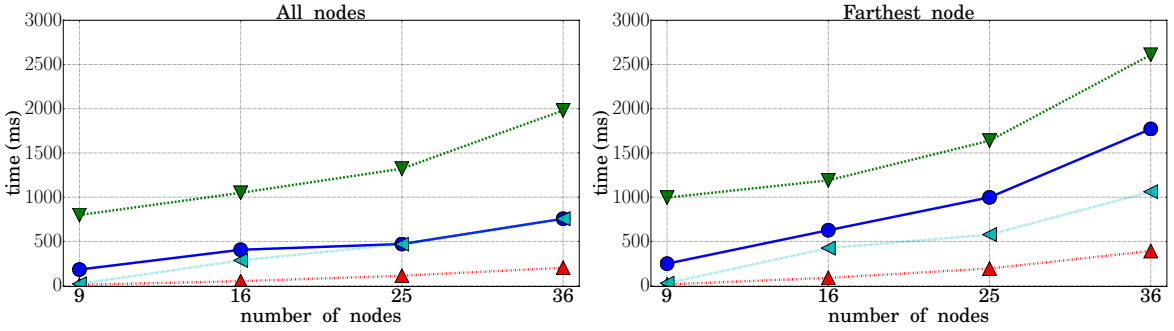
Given the critical nature of some applications that transmit their data through the Smart Grid NAN, it have been considered of importance the analysis of the percentiles of the transit time in conjunction with the average value. In addition, given the size of these networks, it has been considered important to offer, along with the average values for all the network nodes, the values obtained for the farthest smart meter from the concentrator, since these are the values that should be taken into account when planning the network. Figure 4.8 compares the value of the average value and 95th percentile for the HWMP and EA-HWMP cases. As can be seen, these values are always lower (for all network sizes and for all types of traffic) when using the techniques proposed in this chapter. In this figure, a significant improvement is observed in the average and percentile values obtained. In the farthest node case, the improvement is really substantial, with an important reduction of the percentile values for large network sizes.



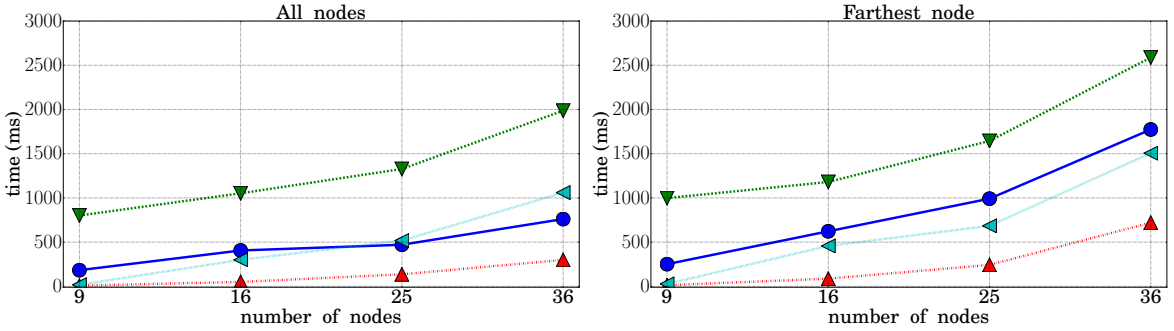
(a) Traffic type 1.



(b) Traffic type 2.



(c) Traffic type 3.



(d) Traffic type 4.

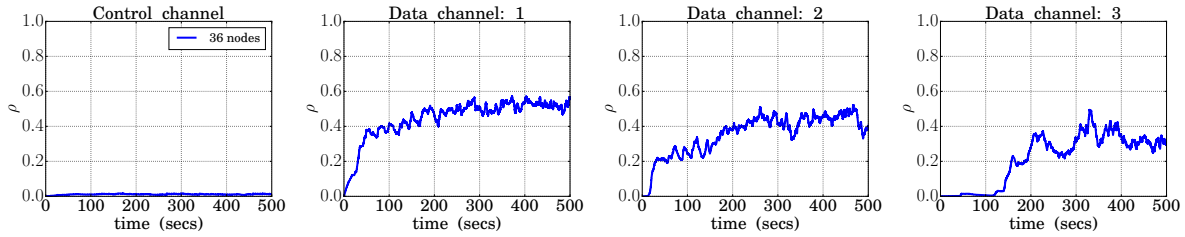
FIGURE 4.8: Network transit time.

Channel utilization factor

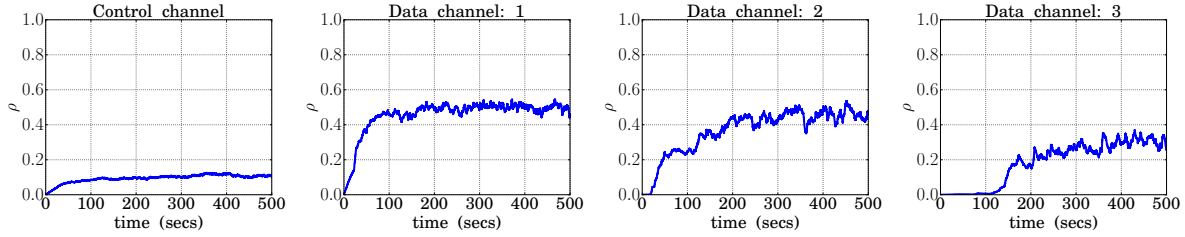
The possibility of having multiple channels implies an increase in the number of successfully transmitted packets to the destination. EA-HMWP assigns the transmission of frames to the

next available channel when the following situations are true. First, the probability of transmitting the frame is calculated taking into account the congestion control functions defined for each traffic type. Second, if the frame will be effectively transmitted, the transmission is assigned to the first available data channel (channel 1 by default). However, if the ρ of this channel is above the accepted threshold for that channel, the transmission will be assigned to channel 2 and so on up to a maximum limit of three data channels.

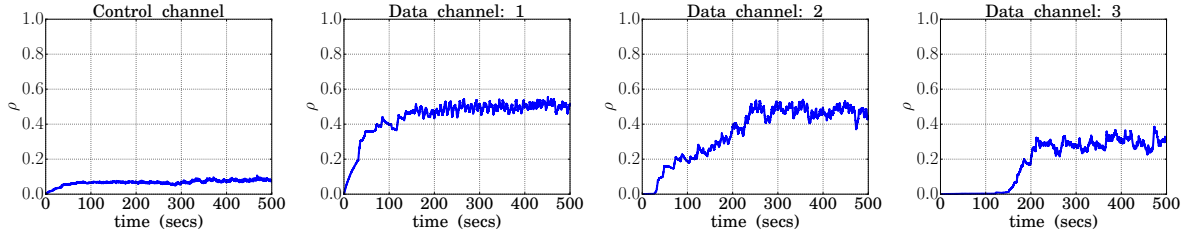
Figure 4.9 shows the utilization factor for the four available channels: three data channels, as mentioned above, plus one control channel. The control channel is used exclusively for the transmission of routing and emergency signaling packets, as will be explained in the next section, and so its utilization factor is very low. However, the control channel utilization factor is greater when the interval time between two successive PREQs is reduced (Figures 4.9a, 4.9b and 4.9c). Regarding the data channels, it can be observed that the ρ value does not achieve high values (it is always below the configured threshold), and therefore, the channels are not congested. Also, the channel utilization decreases from data channel 1 to data channel 3, in concordance with what has been explained in the previous paragraph. On the other hand, Figure 4.10a shows the number of data channels used (n) to transmit the data as the simulation time increases, while Figure 4.10b shows the probability that one or more channels are employed. It can be observed in both figures how, as the network size is increased, it becomes more congested, and therefore a greater number of data channels is needed to transmit the NAN applications.



(a) On-demand mode. Reactive routing information lifetime: 5.12 s.

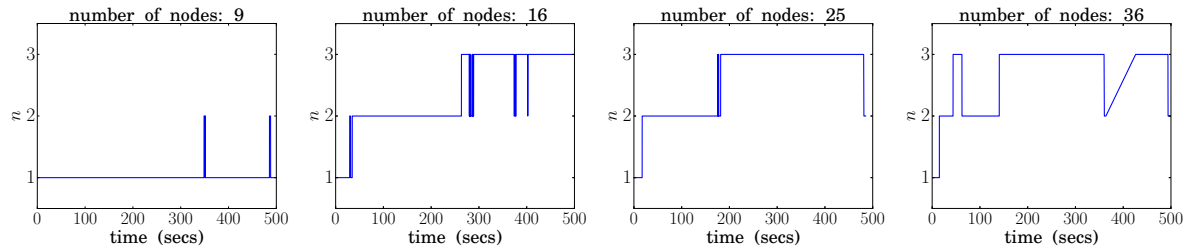


(b) On-demand mode. Reactive routing information lifetime: 0.512 s.

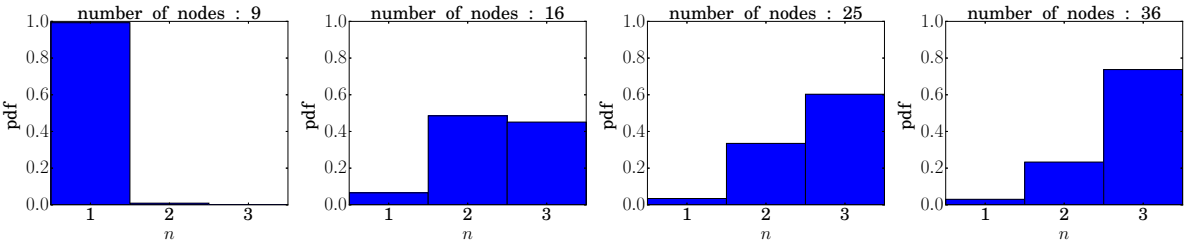


(c) Proactive mode. Proactive routing information lifetime: 5.12 s.

FIGURE 4.9: Channel utilization factor.



(a) channel occupancy (temporary evolution).

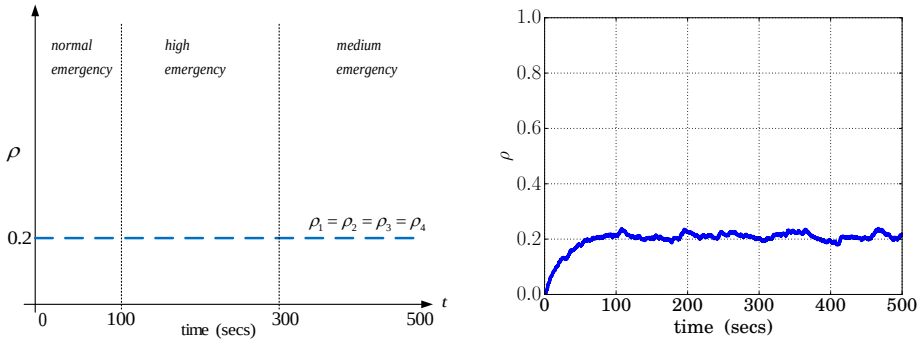


(b) channel occupancy (pdf).

FIGURE 4.10: Channel occupancy for the EA-HWMP. On-demand mode. Reactive routing information lifetime: 5.12 s.

4.3.2.3 Emergency scenario

In the present scenario, the channel allocation and traffic differentiation techniques, congestion control mechanism and the emergency system are implemented and evaluated together. For this purpose, some modifications were made to the previous scenario. First, NAN applications were configured to generate the same amount of data for all types of traffic (Figure 4.11). Second, four emergency sub-scenarios are considered (normal, medium, high and a combination of the previous three). The purpose it to evaluate how the congestion control functions are adapted to the different emergency situations. That is, to give a greater probability of transmission to the most important traffics, specially in emergency situations. These situations are set up by using broadcast messages. For the first three sub-scenarios, at the beginning of the simulation an emergency message is sent by a node to indicate the type of emergency situation throughout the network. On the other hand, in the last sub-scenario (combined) three packets are sent at different times (Figure 4.11a). At the beginning of the simulation (t=0) a normal emergency situation message is transmitted, after 100 seconds a medium emergency situation message is sent, and then after 300 seconds a message of medium emergency situation is sent. The results are evaluated with the same figures of merit (packet delivery ratio, throughput, transit time and channels occupation) and tools used in the previous scenario.



(a) target ρ.

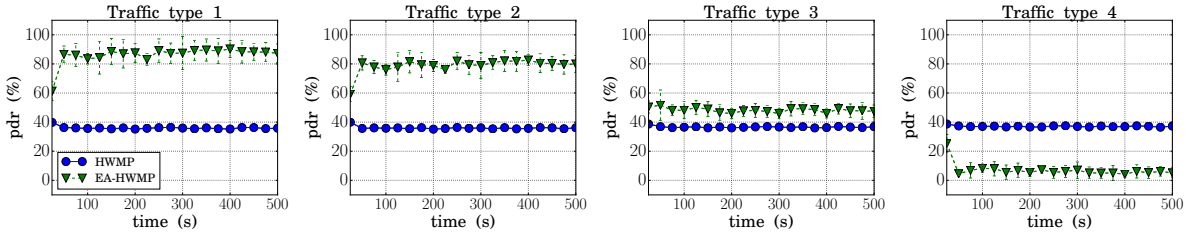
(b) measured ρ.

FIGURE 4.11: Traffic generation for the different NAN applications for the emergency scenarios.

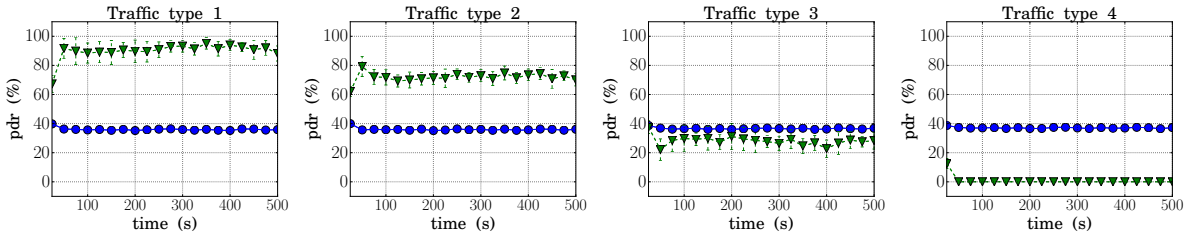
Packet delivery ratio and network throughput

In this subsection, the benefits obtained in terms of packet delivery ratio and throughput are presented in the same two ways as the previous scenario but with some differences. First, HWMP and EA-HWMP are compared taking into account the temporary evolution of the PDR and throughput. Second, the global average values are analyzed.

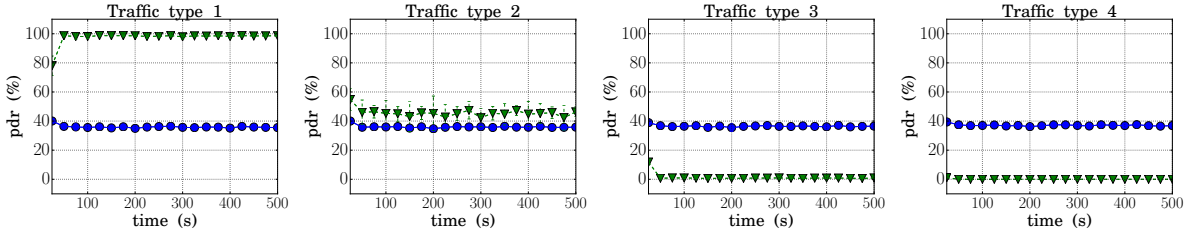
Figures 4.12 and 4.13 show the PDR and throughput evolution over time for the four emergency scenarios and the four traffic types. For the first three scenarios, the same results are highlighted. The PDR and throughput are increased for the higher priority traffic, and these values are increased even more when the emergency situation is high. On the other hand, in situations of high emergency, lower priority NAN applications are discarded, and therefore the PDR and throughput are reduced. In the combined case, as mentioned, the scenario goes from a normal emergency situation (0 to 100 s), to a high (100 to 300 s) and ends in a medium emergency situation (300 to 500 s). The changes of emergency situations can be clearly seen in Figures 4.12d and 4.13d. In a high emergency situation, the system adapts the congestion control functions in order to give a greater probability to NAN applications grouped as traffic type 1. This can be seen with the increase of PDR and throughput from 100 to 200 s. However, the price to pay is a fall of these parameters for the less priority traffic. In the last stage, the PDR and throughput increase when going from high to medium emergency situation.



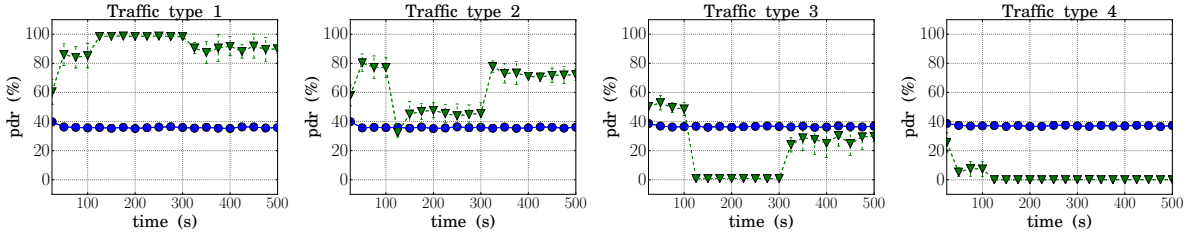
(a) Emergency situation: normal.



(b) Emergency situation: medium.

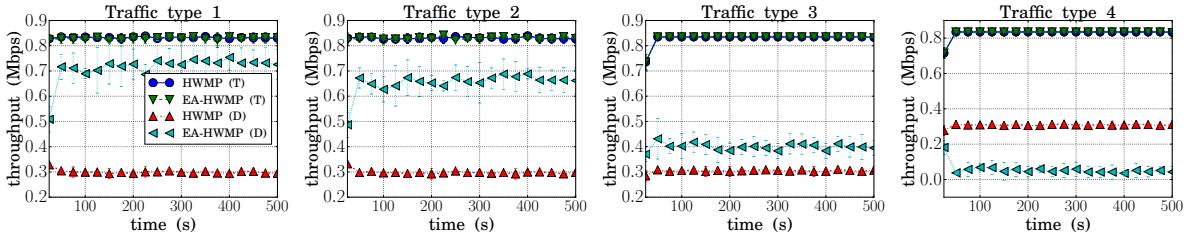


(c) Emergency situation: high.

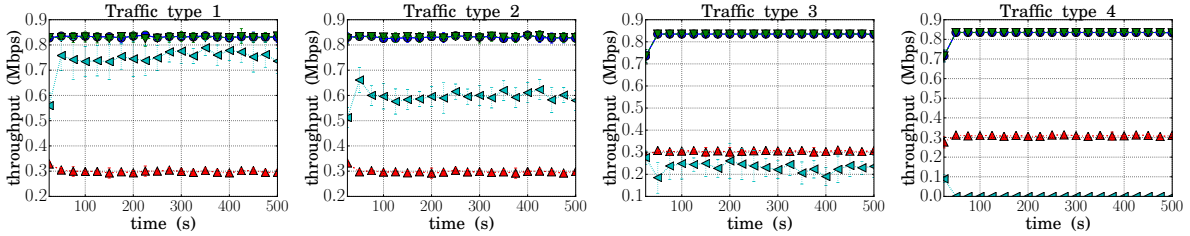


(d) Emergency situation: combined.

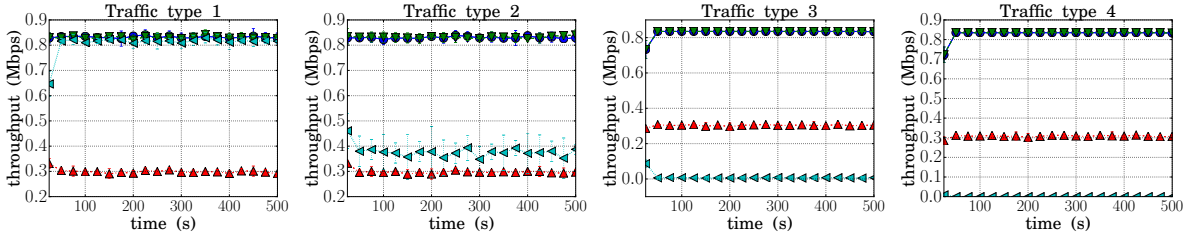
FIGURE 4.12: Packet delivery ratio evolution over the time for different emergency situations (network size: 25).



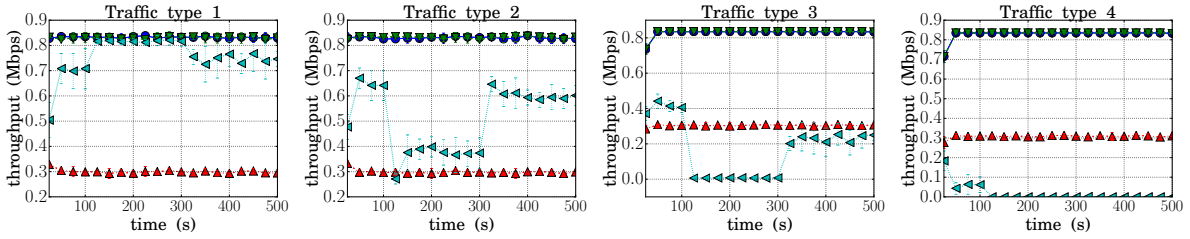
(a) Emergency situation: normal.



(b) Emergency situation: medium.



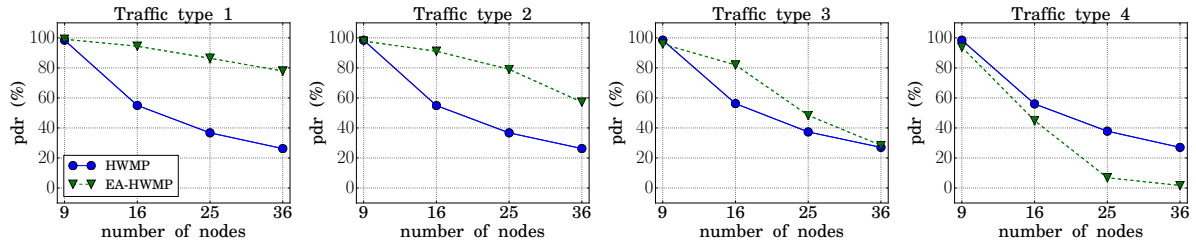
(c) Emergency situation: high.



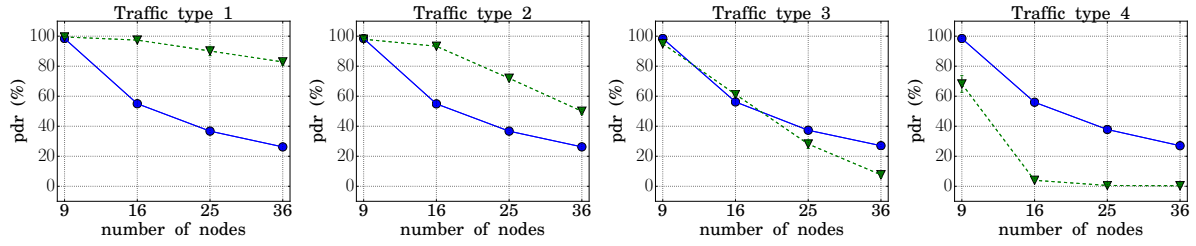
(d) Emergency situation: combined.

FIGURE 4.13: Network throughput evolution over the time (network size: 25).

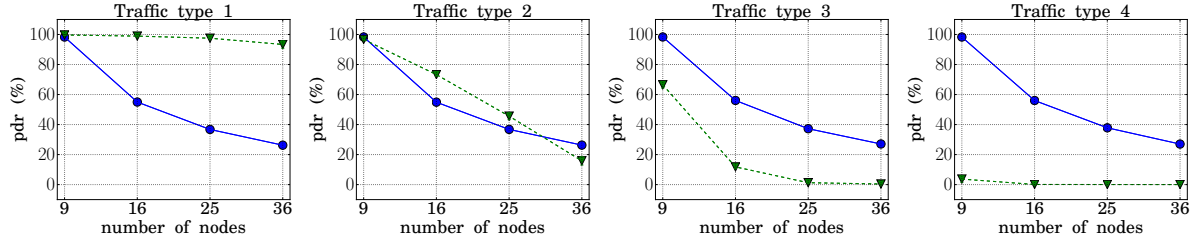
The PDR and throughput average values (together with the 95% confidence interval) are shown in Figures 4.14 and 4.15, for different network sizes. As the network size is increased, it becomes more congested and so the PDR decreases for all traffics (Figure 4.14). However, it is important to highlight some important aspects. On the one hand, this decrease is smaller with EA-HWMP for priority traffic types. In addition, emergency situations give a greater probability of transmission to the most important traffic. Therefore, the PDR is increased from the normal emergency situation to the high one for the traffic type 1. Meanwhile, for the less priority traffic types, the PDR is reduced by discarding packets of some NAN applications. On the other hand, Figure 4.15 complements the result obtained by the PDR parameter. That is, the network throughput is higher for traffic type 1, and it is even greater for the high emergency situation. For the less priority traffic types, the throughput decreases when the emergency situation changes from normal to high.



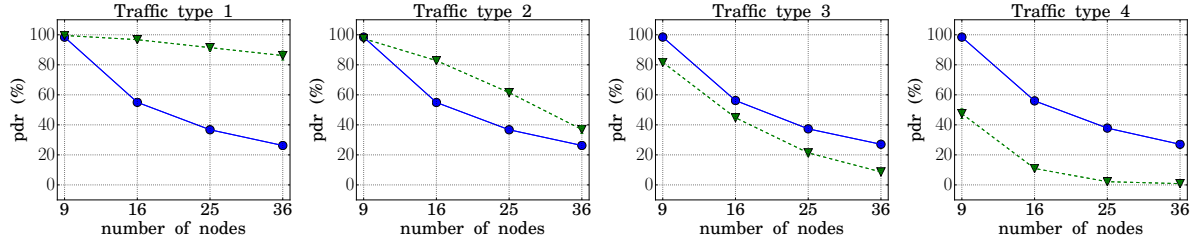
(a) Emergency situation: normal.



(b) Emergency situation: medium.

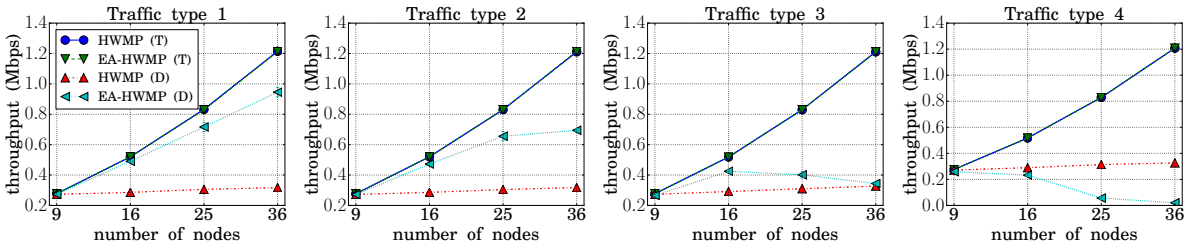


(c) Emergency situation: high.

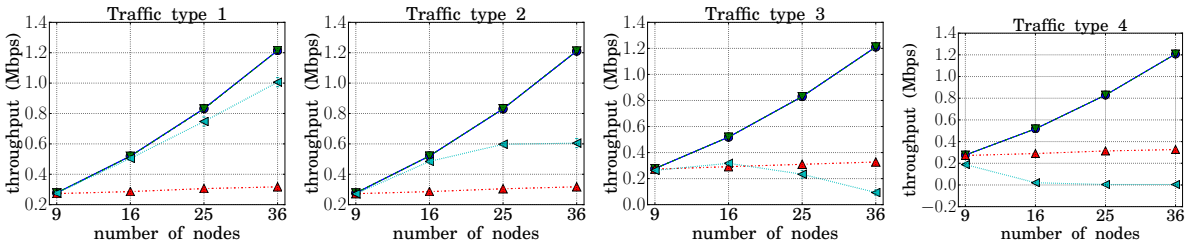


(d) Emergency situation: combined.

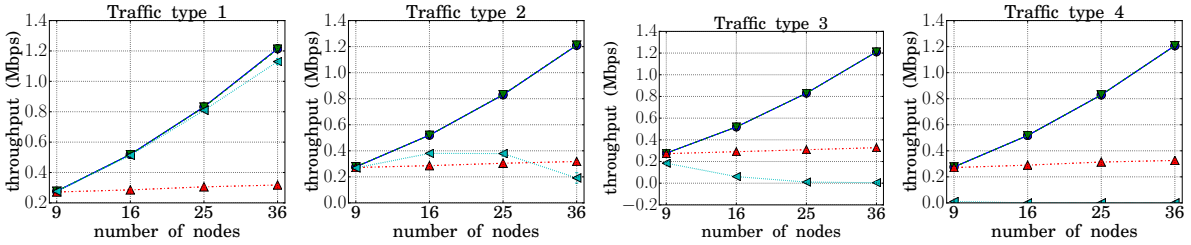
FIGURE 4.14: Packet delivery ratio for different emergency situations.



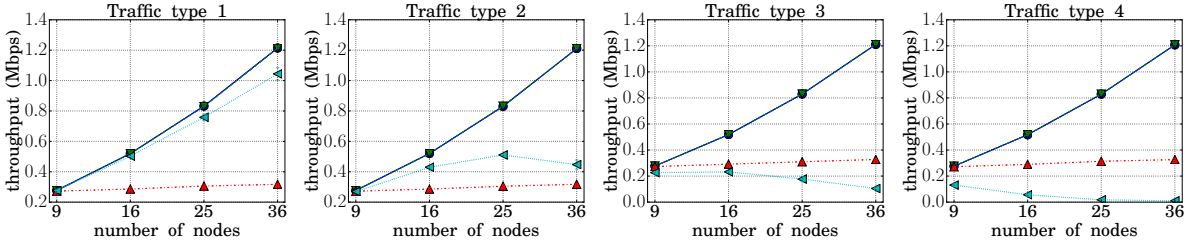
(a) Emergency situation: normal.



(b) Emergency situation: medium.



(c) Emergency situation: high.

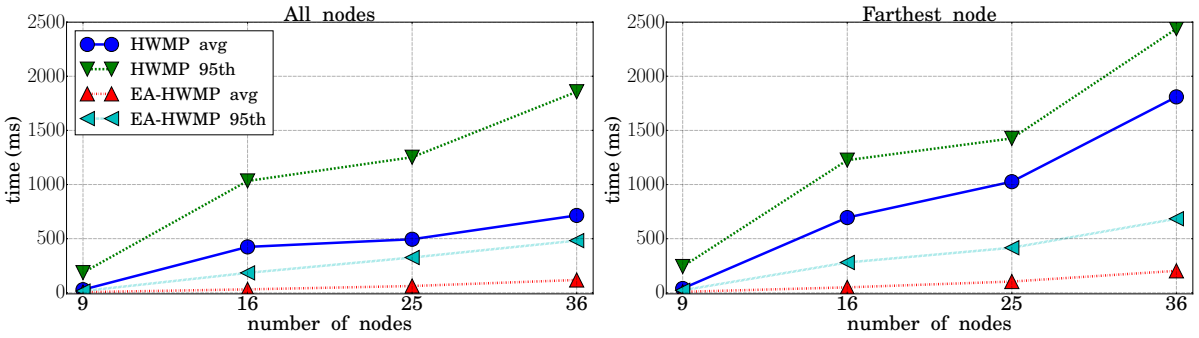


(d) Emergency situation: combined.

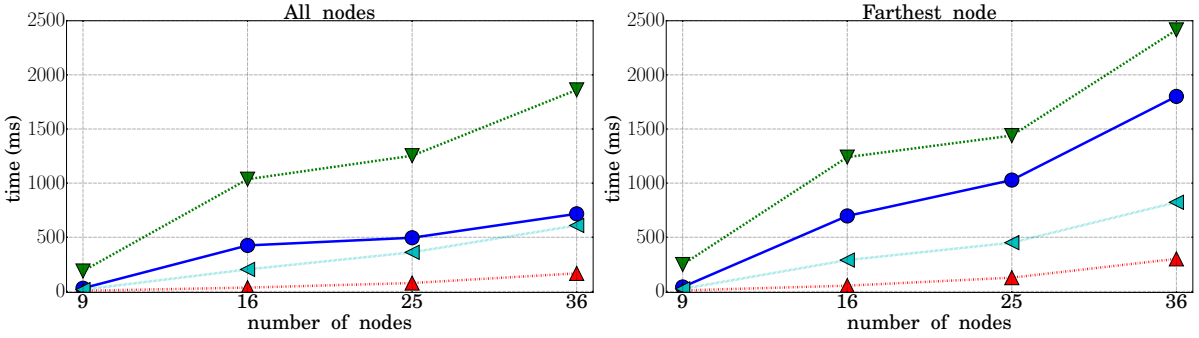
FIGURE 4.15: Network throughput for different emergency situations.

Network transit time

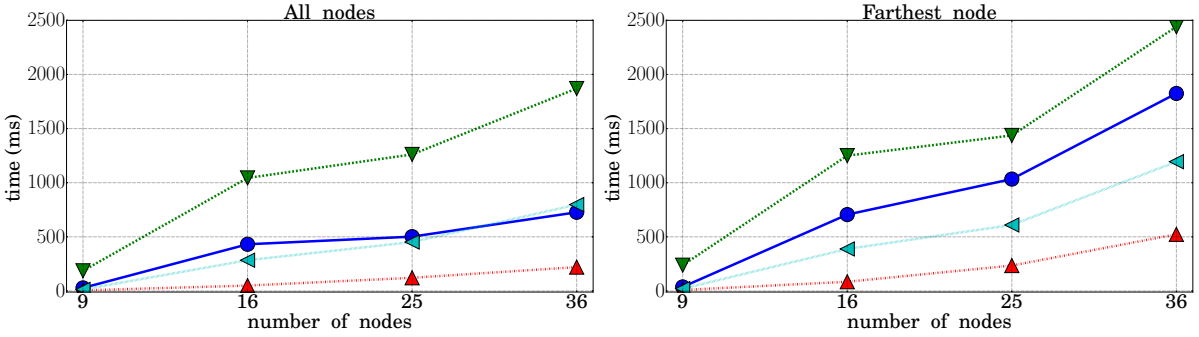
In the same way as the previous scenario, the analysis of the percentiles of the transit time, together with the average values are presented, both averaged for all the network nodes, and taking into account only the node farthest from the concentrator. Figure 4.16 shows the results for the combined scenario. As can be seen, these values are always lower (for all network sizes and for all types of traffic) when using the techniques proposed in this chapter. Similarly, a significant improvement is observed in the average and percentile values obtained in the farthest node case.



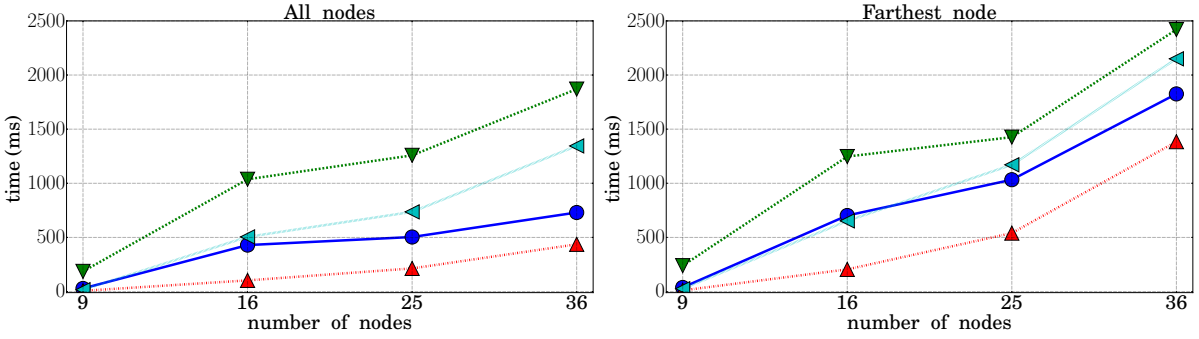
(a) Traffic type 1.



(b) Traffic type 2.



(c) Traffic type 3.



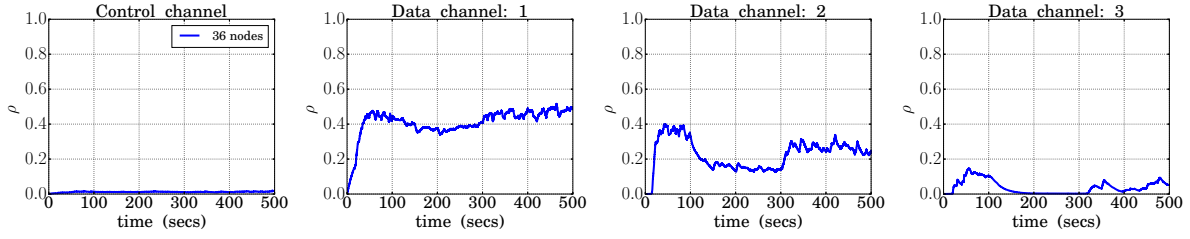
(d) Traffic type 4.

FIGURE 4.16: Network transit time (Emergency situation: combined).

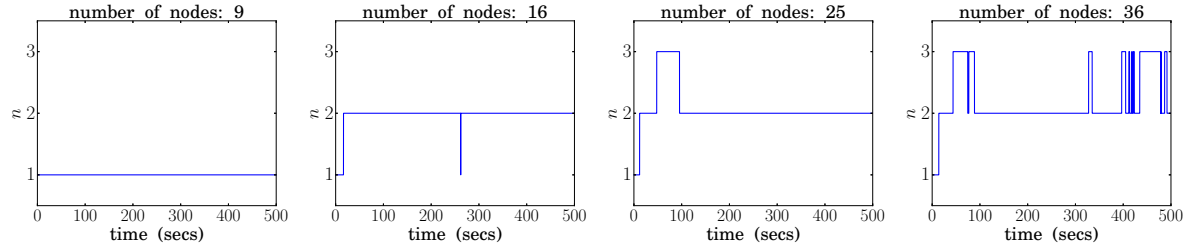
Channel utilization factor

Figure 4.17a shows the channel utilization factor measurement for the EA-HMMP protocol, for the nearest smart meter node to the concentrator, and for the four available channels (one

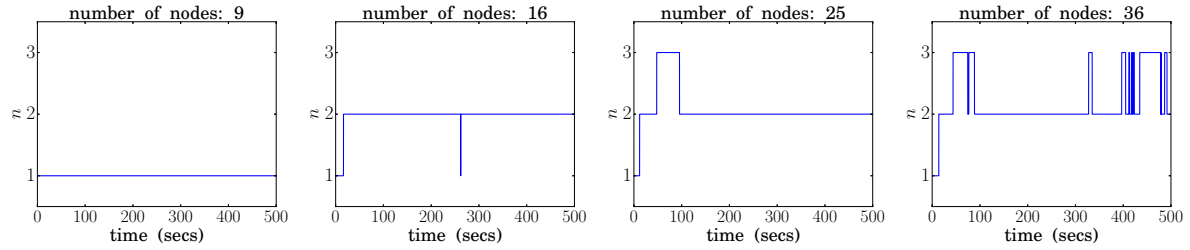
control channel and three data channels). As in the previous scenario, it can be seen how the control channel, dedicated to transmit only emergency and routing control messages, is never congested, since its ρ value is very low. Also, data channels are not congested, since the value of ρ never reaches high values (always below the ρ threshold). On the other hand, the channel utilization decrease from the data channel 1 to the data channel 3, and in particular, data channel 2 presents the fall of the value of ρ (100 to 300s) when the system is in a high emergency situation. Finally, Figures 4.17b and 4.17c show the results for the channel occupancy in terms of temporary evolution and probability. In the same way as the previous scenario, it can be seen how as the network size is increased, a greater number of channels are used to transmit the applications.



(a) ρ measured.



(b) channel occupancy (temporary evolution).



(c) channel occupancy (pdf).

FIGURE 4.17: Channel occupancy for the EA-HWMP (emergency situation: combined).

Overall, this first solution has shown the benefits obtained when a congestion control mechanism, multichannel allocation scheme and an emergency system are applied together to deal with high network loads. In the next section, a new fair and distributed congestion control mechanism is proposed, implemented and evaluated for NANs.

4.4 A Fair and Distributed Congestion Control Mechanism for Smart Grid Neighborhood Area Networks

As previously said, the second congestion control control proposal deals mainly with the provision of fairness of the distribution of the network resources. Note that some network nodes

might monopolize the use of the channels due to their higher traffic generation rate, or even their geographical position. To this end, a distributed solution that combines several algorithms is proposed, implemented and evaluated where traffic differentiation is also taking into account.

Unlike the strategies presented previously, the congestion control mechanism is now applied in the source nodes. Thus, depending on the sensed degree of network congestion, source nodes may increase or decrease the rate at which they generate their packets. This way, the mechanism avoids unnecessary transmissions of packets that will be likely discarded later on their way to their destination. To this end, new signaling messages must be transmitted from relay nodes to source nodes.

The proposed solution does not rely on a specific network, mac or even a physical layer. Therefore, the distributed solution is now evaluated in the context of a wireless ad hoc network composed by a set of Smart Grid meter devices where the AODV routing protocol and IEEE 802.11 ac physical layer are used for the network and physical layer respectively.

4.4.1 Proposed solution

Table 4.6 presents the principal notation of the parameters used to implement the distributed solution. Remember that a NAN comprises a set of N nodes corresponding to Smart Meters (SM) plus a data concentrator node (aka sink) that acts as a gateway to the WAN segment. Smart meters collect information from their HAN segment (e.g., gas meter, water meter, electricity meter, other sensors) and exchange them with the data concentrator node directly or through other smart meters that serve as relay nodes to form a multi-hop communication network. Besides, the nodes of the smart grid communicate using wireless transmissions. More precisely, we assume that the nodes implement the IEEE 802.11 standard. Furthermore, traffic flows have typically different QoS requirements in terms of throughput and network transit time. More specifically, it is assumed again that flows are divided into K different traffic types, numbered 1,...,K, where traffic type 1 has the most stringent requirements. Each flow generates data at a given rate that it is referred to as its nominal rate. This data rate is determined by the smart grid application and does not necessarily comply with the available communication resources of the NAN. x_n^k has been used to denote the nominal rate (in bps) of the flow originating at node n ($n = 1, ..., N$) of traffic type k ($k = 1, ..., K$) where N refers to the number of nodes. Finally, the number of source nodes is indicated by S where $S < N$.

In previous solutions, several metrics have been used to evaluate the performance of the NAN as well as those attained by the flows present in the NAN. In this new solution, new metrics for network performance evaluation have been included and modified:

- The *averaged channel utilization factor* on each node have been considered. A modification for the ρ measurement has been made, and more precisely, ρ_n refers to the occupation of the radio channel at node n averaged over a period of T_A seconds.
- In the analysis it has been included the *Jain's index* [90] calculated over the attained throughput for flows belonging to the same traffic type but originating from different sources. The Jain's index is a common metric to assess the level of fairness between N entities. Its value ranges from $1/N$ (worst case) to 1 (best case). In this case, it is maximum when traffic flows attain the same throughput, regardless of their distance to the concentrator node or their particular packet generation rate.

TABLE 4.6: Principal notation.

Parameter	Description	Units
	System parameters	
N	Number of nodes	
K	Number of traffic types	
S	Number of source nodes	
x_n^k	Nominal rate of the flow originating at node n of traffic type k	bps
ρ_n	Current utilization of the radio channel at node n	
	Proposed solution parameters	
ρ_U	Over-utilization threshold for the radio channel	
ρ_L	Under-utilization threshold for the radio channel	
T_A	Periodicity of Algorithms 8 and 10	sec
y_j^k	Estimated rate for the flow originating at node j of traffic type k	bps
β_I	Increasing factor in case of under-utilization	
β_D	Decreasing factor in case of over-utilization	
$z_{j,n}^k$	Adapted rate returned by node n for the flow originating at node j of traffic type k	bps
δ^k	Unitary increment for traffic type k	bps
α^k	Percentage (weight) assigned to traffic type k	
s_j^k	Implemented flow rate for the flow originating at node j of traffic type k	bps

The proposed FDCC (Fair Distributed Congestion Control) solution comprises 3 algorithms and is distributed in the sense that it operates both at source and relay nodes. Bear in mind that a node (with the exception of the data concentrator) can be a relay, a source or both. The rationale of our algorithms is to enable high utilization of the radio channel while avoiding the occurrence of long-lasting congestions causing packet losses and collisions.

To achieve this trade-off, the solution assigns a certain portion of the (estimated) available communication resources to each flow based on their traffic type and source. Note that the size of these portions are easily parameterized. FDCC breaks down as follows. Through Algorithm 8, relay nodes monitor their current utilization of the radio channel as well as the current rate for each flow hopping through them. Depending on the current value of radio channel utilization, Algorithm 9 is triggered, and delivers in return adapted source rates for each flow. Lastly, newly computed source rates are notified to the associated source nodes that readjust their flow rates, see Algorithm 10.

Note that these latter notifications are the only control packets generated by our algorithms. Consequently, during each computation period (e.g., one second in our experiments), the solution generates at most a number of control packets equal to the number of relay nodes multiplied by the number of sources. Besides, the algorithms imply only a few instructions and their requirements in memory and CPU are low.

These three algorithms are described in greater detailed in the following subsections.

4.4.1.1 Detecting over- and under-utilization of the radio channel

Algorithm 8 is implemented at each relay node and monitors if the radio channel is over-utilized, or alternately, under-utilized. In practice, for each relay node n, the algorithm verifies

if $\rho_L < \rho_n < \rho_U$ where ρ_L and ρ_U are constant values in between 0 and 1. Additionally, through Algorithm 8, relay nodes evaluate the current rate for each flow hopping through them. It is used y_j^k to denote the current rate calculated for the flow of traffic type k and originating from node j. Note that y_j^k is computed as a simple arithmetic average over a period of time A and that, in general, y_j^k and x_j^k (nominal rate) differ.

If $\rho_n \leq \rho_L$ or $\rho_n \geq \rho_U$, Algorithm 8 triggers the readjustment of rate for the flows hopping through node n in an attempt to get the radio channel utilization ρ_n back in the interval $[\rho_L; \rho_U]$. More precisely, in case of over-utilization (resp. under-utilization), Algorithm 8 calls Algorithm 9 with the current values of y_j^k and a parameter β_D (resp. β_I) which is a constant between 0 and 1 (resp. larger than 1). Otherwise, if ρ_n, lies in between ρ_L and ρ_U, then Algorithm 8 remains idle.

Algorithm 8: Run at each relay node n to detect under- or over-utilization of the radio channel.

Input: T_A, ρ_L, ρ_U, β_D, β_I

```
1 while True do
2     Update value of ρ_n and y_j^k
3     if ρ_n ≥ ρ_U then
4         Throttle flows: compute z_j^k following Algorithm 9 (y_j^k, β_D)
5         Notify the sources with the new values of z_{j,n}^k
6     else if ρ_n ≤ ρ_L then
7         Relax flows: compute z_j^k following Algorithm 9 (y_j^k, β_I)
8         Notify the sources with the new values of z_{j,n}^k
9     Wait T_A seconds
```

4.4.1.2 Determining the adapted source rates of flows

Algorithm 9 is also implemented on the relay nodes. It computes the adapted source rates for each flow that will then be notified to the sources. Let $z_{j,n}^k$ denote the adapted source rates returned by node n for the flow originating at node j of traffic type k. With FDCC the values of $z_{j,n}^k$ strongly depend on the current assessment of the radio channel utilization. This algorithm is called by Algorithm 8 either to increase (with proportion β_I), or alternately, to decrease (with proportion β_D) the rates from the sources. Nonetheless, flows from different traffic types may undergo different adaptation as described below.

Initially, the adapted rate that will be assigned by node n to a flow from source j with priority k is set to 0, namely $z_{j,n}^k = 0$. The algorithm is composed of two iterative steps.

In the first step (lines $4-7$), the algorithm aims at allocating a certain portion of the available resources to each flow according to their current rate and traffic type. Their respective rates are increased in a while loop with increments of length δ_k. For a given flow of traffic type k, it stops increasing when a percentage, denoted by α_k, of the current rate has been reached, namely $z_{j,n}^k < \alpha_k \cdot y_j^k$.

In a second step (lines $8-11$), if the threshold on the overall transmission rate has not been reached ($\sum z_{j,n}^k < F \cdot \sum y_j^k$ where F is an input parameter of Algorithm 9 that sets the size of increments or decrements -β_I or β_D from Algorithm 8- for the overall transmission rate),

the resources left are shared among the traffic types and sources with the same principle, i.e. increasing with increment δ_k.

Both thresholds α_k and δ_k depend on the flows traffic type k. Typically, we have $\alpha^1 > \alpha^2 > \cdots > \alpha^K$ and $\delta^1 \geq \delta^2 \geq \cdots \geq \delta^K$. The parameter setting will be discuss in further detail in the next section.

Algorithm 9: Run at each relay node n to determine the adapted source rates of flows.

Input: y_j^k, F
Output: $z_{j,n}^k$

$z_{j,n}^k = 0$
while $\sum_{j,k} z_{j,n}^k < F \cdot \sum_{j,k} y_j^k$ **do**
  for (j,k) *in* $\{1,\ldots,S\} \times \{1,\ldots,K\}$ **do**
    if $z_{j,n}^k < \alpha_k \cdot y_j^k$ **and** $\sum_{j,k} z_{j,n}^k < F \cdot \sum y_j^k$ **then**
      $z_{j,n}^k = z_{j,n}^k + \delta^k$
while $\sum_{j,k} z_{j,n}^k < F \cdot \sum_{j,k} y_j^k$ **do**
  for (j,k) *in* $\{1,\ldots,S\} \times \{1,\ldots,K\}$ **do**
    if $\sum_{j,k} z_{j,n}^k < F \cdot \sum y_j^k$ **then**
      $z_{j,n}^k = z_{j,n}^k + \delta^k$

4.4.1.3 Implementing the newly adapted flow rates at the sources

Algorithm 10 is implemented at the source nodes. For its flow of traffic type k, the source node j may receive multiple values of $z_{j,n}^k$ where n refers to the relay node that computed the proposed flow rate. Every T_A seconds, Algorithm 10 readjusts the adapted rate for its flow of traffic type k as the minimum of received value $z_{j,n}^k$. Denoting s_j^k as the adjusted rate of the flow of traffic type k originating at node j, we have: $s_j^k = \min_n(z_{j,n}^k)$ where n spans over the relay nodes that returned an adapted rate to node j. Note that in some case the nominal rate of a flow can be found to be less than its readjusted rate (i.e., $x_j^k < s_j^k$). Should it be the case, then the implemented flow rate would simply be the nominal one.

Algorithm 10: Run at each source node j to adjust the adapted flow rates.

Input: T_A, $z_{j,n}^k$
while *True* **do**
  Initialize all $z_{j,n}^k = \infty$
  Collect values $z_{j,n}^k$ from the relay nodes
  Compute $s_j^k = min_n(z_{j,n}^k)$
  for k *in* $\{1, \ldots, K\}$ **do**
    if $y_j^k < s_j^k$ **and** $x_j^k \geq y_j^k$ **then**
      Increase y_j^k to its new value s_j^k
    if $y_j^k > s_j^k$ **then**
      Decrease y_j^k to its new value s_j^k
  Wait T_A seconds

4.4.2 Results and Discussion

Now the experimental setup to evaluate the performance of our proposed FDCC solution is described. First, general information about the relevant parameters values chosen for the simulations is provided. Next, the obtained results for three different scenarios are shown. The first one consists of a tree topology used to explain over a simple scenario the proposed mechanism, while the other two represent more realistic topologies for Smart Grid NANs.

4.4.2.1 Simulation details

FDCC is also implemented in C++ within the ns-3 [2] simulator. The wireless communications between nodes are operated using the recent amendment 802.11ac of the IEEE 802.11 standard on a single 20MHz channel with a MCS (Modulation Coding Scheme) set to 0 so that transmissions occur at a physical rate of 6.5 Mbps. The maximum number of packets that can be buffered in every node has been set to 100 packets. The different traffic types are presented in Table 4.7 where now the packet length and the packet generation rate distributions have been selected as exponential. The exponential distribution has been used in order to have a vast number of NAN applications.

Paths between source and destination nodes are found using the AODV [61] routing protocol wherein the weight of links derives from the expected number of frames transmitted to send a packet (ETX). The ns-3 simulator includes by default the AODV routing protocol with the hop count metric, the ETX metric implemented by [91] has been used. Anyway, as it was mentioned in previous sections, FDCC is agnostic to the routing protocol, and so other routing protocols can be used.

Each simulation corresponds to 100 seconds of network operation, and is run several times with different pseudo-random generator seeds. Besides, the first twenty seconds of each simulation are considered as a transient regime, and so results are not gathered during this time. Table 4.8 shows the main simulation parameters.

FDCC solution involves several parameters that must be set before its execution. Guidelines on how to select them are briefly presented. K is simply determined by the number of traffic types. T_A sets the frequency at which the solution results are updated. A value near 1 sec seems to be good trade-off between reactiveness and bandwidth usage. β_D (resp. β_I) specifies the level at which sources decreases (resp. increases) the flow rates in case of over-utilization (resp. under-utilization). It has been selected β_D =1.05 and β_I =0.75 in order to gradually coming closer the saturation threshold while quickly going away from it if reached. ρ_U and ρ_L defines the interval at which the radio channel is deemed adequately utilized (not under- nor over-utilized). In the case of links operating over IEEE 802.11, having an utilization factor between 0.7 and 0.8 provides high but not saturated utilization of the resources. The values of α^i indicate the relative importance of each traffic type with respect to the others. It is recommended taking values in the range of 0.5 and 0.8. Finally, δ^i's values set the granularity of the solution and as so must be set small enough. In this case, a value of 1 was selected. Table 4.9 summarizes the parameter values used for the congestion control mechanism.

For the performance evaluation, FDCC and the EDCA mechanism [55] are compared. As mentioned in Chapter 2, in EDCA the traffic is classified in different categories according to their traffic types. Each node has several internal queues with different parameters to prioritize traffic according to their categories. The ns-3 default parameters, also used in the simulations

for EDCA, are shown in Table 4.10. Recall that the arbitration interframe space (AIFS) is the interval of time that the station must sense the medium as idle before trying to transmit the next frame. On the other hand, the transmission opportunity (TXOP) represents an allocation of time that can be used to transmit one or several frames in a sequence. Finally, for each category the minimum and maximum contention window (CW) are the limits from which the random backoff is uniformly chosen. Keep in mind that the EDCA mechanism is local. It processes frames according only to their traffic types and does not differentiate the flows from the different sources. Furthermore, EDCA mechanism is also independent of current state of the network (e.g., congestion).

TABLE 4.7: NAN applications transmitted over the Smart Grid.

Traffic type	**Applications**	**Packet length average (Bytes) and PDF**	**Packet generation rate (per sec) average and PDF**
1	Demand Response, Outage Management	200 Exponential	100 Exponential
2	Video surveillance, Overhead Transmission Line Monitoring, Substation Automation systems (SASs)	200 Exponential	100 Exponential
3	Home Energy Managament (HEM), Electric Vehicles (EVs) Charging	200 Exponential	100 Exponential
4	Meter Data Management	200 Exponential	100 Exponential

TABLE 4.8: Main simulation parameters.

Description	**Value**
Network simulator.	ns-3.28
Simulation time	100 s
Tranport layer	UDP
Random number generator	MRG32k3a
Routing protocol	AODV
Routing metric	ETX
Maximum queue size	100 packets
Wireless physical layer	802.11ac
Modulation coding scheme	0
Short guard interval	1
Frame aggregation factor	0
Channel width	20 MHz

TABLE 4.9: FDCC used parameters.

Parameter	Value
K	4
T_A	1 sec
β_D	0.75
β_I	1.05
ρ_U	0.8
ρ_L	0.7
$[\alpha^1, \alpha^2, \alpha^3, \alpha^4]$	$[0.8, 0.7, 0.6, 0.5]$
$[\delta^1, \delta^2, \delta^3, \delta^4]$	$[1, 1, 1, 1]$

TABLE 4.10: EDCA used parameters.

Traffic type	**CWmin**	**CWmax**	**AIFS**	**Max TXOP**
4	15	1023	7	0
3	15	1023	3	0
2	7	15	2	3.008 ms
1	3	7	2	1.504 ms

4.4.2.2 Tree topology scenario

In the first scenario, a tree topology made of six nodes with three traffic sources (S1, S2 and S3), two relays (R1 and R2) and one destination (D) is considered as depicted in Figure 4.18. As discussed above, each source node delivers traffic belonging to four different traffic types. In any case, packets are first sent to node R1, then forwarded to node R2 before finally reaching node D. The distances between nodes are chosen so that, together with the selected channel model, they lead to the network topology shown in the figure. Note that with a packet generation rate at each source node of 400 pkt/s (viz 0.64 Mbps), the combined workload of S1, S2 and S3 amounts to 1.92 Mbps and will undoubtedly overflow node R1, whose transmission rate of 6.5 Mbps is reduced by half due to DCF overheads and it is shared with its four neighbors.

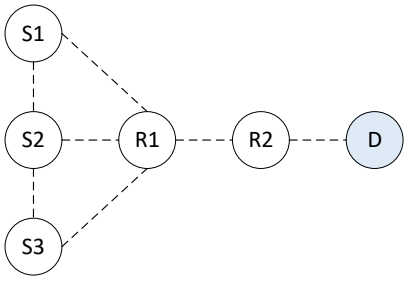


FIGURE 4.18: Tree topology scenario.

Channel utilization factor and buffer occupancy

The performance results obtained with the proposed mechanism (Fair Distributed Congestion Control, FDCC) are shown. These results are compared with the obtained without congestion

control (Uncontrolled), and with those obtained when the Enhanced Distributed Coordination Funtion (EDCA) is activated. To begin with, the channel utilization factor at the relay node R1 is anayzed. Channel utilization here refers to the occupation factor of the radio channel averaged over a period of time of 1 second. Figure 4.19 shows the corresponding results. The blue line (labelled *Uncontrolled*) represents the case where FDCC does not operate and sources are free to generate packets at their initial rate, namely 100 pkt/s for each traffic type. The channel utilization factor nears 0.8 (meaning that R1 perceives the radio channel as busy about 80% of the time). In practice, the radio channel can be viewed as fully saturated since the 20% left corresponds to unavoidable delays such as those incurred by the contention window. On the other hand, the red line (labelled *EDCA*) represents the case where the proposed solution does not operate but the EDCA mechanism does. Here, each traffic type is assigned to an specific access category according to its priority (see Table 4.7). In this case, the channel utilization factor nears 0.84 (meaning that R1 perceives the radio channel as busy about 84% of the time). Finally, the green line corresponds to the case where the proposed solution is applied (labelled *FDCC*). The ρ factor oscillates between 0.6 and 0.8 with a mean value of 0.71. This behavior indicates that, with FDCC, channel saturations may occur but they are rapidly vanishing. On the other hand, Figure 4.20 shows the buffer occupancy in the same relay node R1. It can be seen that, when no mechanism is applied, the buffer is completely full most of the time, resulting in uncontrolled packet losses. In a similar way, when EDCA is activated, the buffer occupation is lower, but the buffer is also filled numerous times, and therefore packet losses also happen. Finally, when using FDCC, the controlled rate reduction makes the occupation of the buffer much lower, not reaching the total occupancy at any time and thus avoiding packet losses. The mean values for the Uncontrolled, EDCA and FDCC cases are 93.55, 19.88 and 9.56 packets respectively.

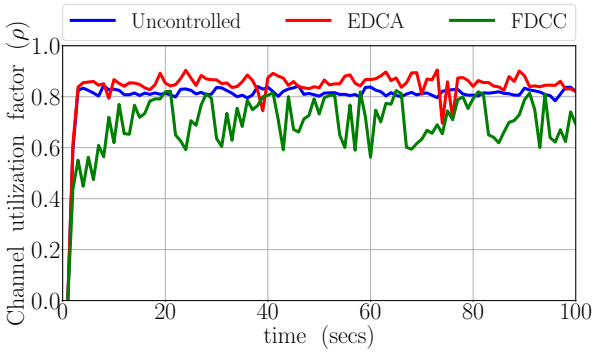


FIGURE 4.19: Evolution of the channel utilization factor seen by node R1.

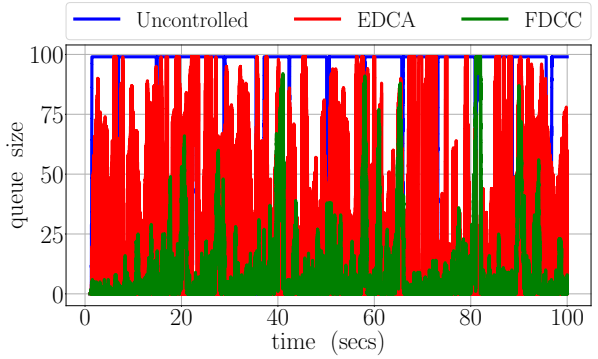


FIGURE 4.20: Evolution of the buffer occupation at node R1.

Packet generation rate

The packet generation rates at the source node S1 are now studied. Figure 4.21 represents, for each traffic type, the minimum and maximum values, together with the percentile boxes (25, 50 and 75-percentiles). As expected, in the absence of a generation rate control mechanism, this rate holds its initial value of 100 pkt/s for any type. On the other hand, with the help of FDCC solution, as can be observed that the traffic types exhibit different patterns. For instance, the traffic type 1 boxplot indicates that the generation rate may vary from 75 up to 95 pkt/s while its median value is approximately 80 pkt/s. Looking at the less critical traffic types, their packet generation rate tends to be significantly smaller but yet far from 0 though.

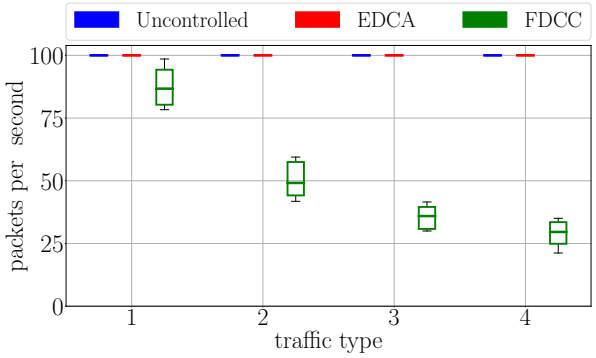
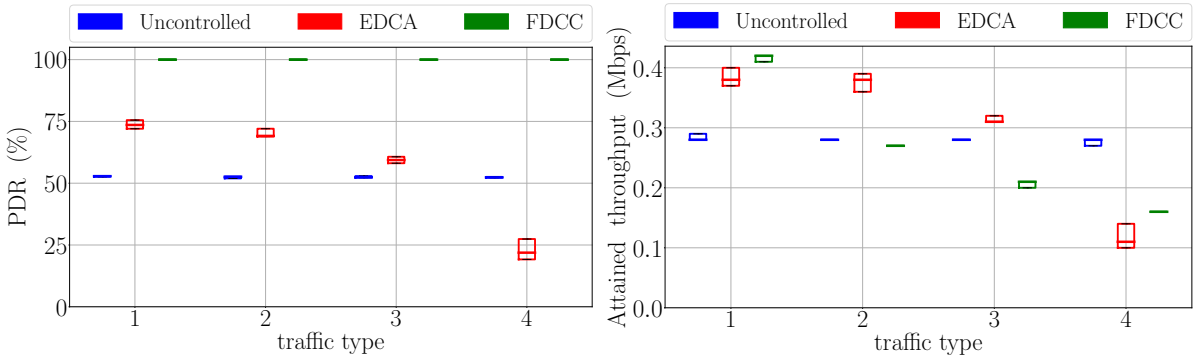


FIGURE 4.21: Packet generation rate at node S1 for each traffic type.

Packet delivery ratio (PDR) and network throughput

In Figure 4.22a, the packet delivery ratio (PDR) attained for each traffic type is presented. Because the radio channel tends to be saturated in the absence of our solution (see Figure 4.19), the corresponding PDR for traffic type 1 only peaks approximately at 50%. Besides, as no distinction is made between the different traffic types, the values is the same for all of them. This fact does not occur when EDCA is applied. As can be seen that in the figure, the PDR value decrease from 75% to 25% depending on the traffic type. Finally, with FDCC, the fact of having a previous generation rate control results in a PDR of 100%.

Combining the two previous performance parameters, generation rate and PDR, the attained throughput can be obtained by the different applications depending on the traffic type they belong to, which is shown in Figure 4.22b. This parameter refers to the number of bytes that were successfully transmitted from the sources to the destination node. Not surprisingly, as can be seen that in the absence of any control mechanism, all traffic types exhibit the same pattern with a constant throughput equal approximately to 0.28 Mbps. On the other hand, both EDCA and FDCC mechanisms cause different attained throughput depending on the traffic type. As will be shown with the next evaluation scenario, another advantage of FDCC is a better distribution of this throughput between all the network nodes.



(a) PDR.

(b) Attained throughput.

FIGURE 4.22: Packet delivery ratio and attained throughput at node D for each traffic type.

Network transit time

The network transit time values are shown in Figure 4.23. Again, when no additional mechanisms are applied, the value is the same for all traffic type. As can be seen, both EDCA and FDCC make the values lower and dependent on the traffic type. Besides, FDCC achieves significant improvements in this parameter for all traffic types. Interestingly, unlike EDCA, this decrease with FDCC is obtained while maintaining the same proportion of traffic from the farthest nodes to the gateway, as will be shown again with the next scenario.

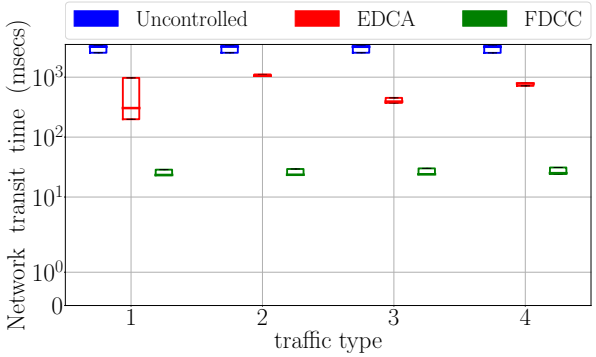


FIGURE 4.23: Network transit time for each traffic type.

Fairness with unbalanced sources

As a final experiment in this simple scenario, the ability of the FDCC mechanism to regulate the rate of all source nodes at the same value and for each traffic type is shown. In this case, the generation rate of the sources S1, S2 and S3 has been configured, for each traffic type, to 300, 200 and 100 pkt/s respectively. Figure 4.24 shows the results obtained once the FDCC mechanism has performed its rate control. As can be seen, for traffic type 1, the rate of sources S1 and S2 has been reduced to the same value of approximately 150 packets per second since the network does not have enough resources to serve the highest quantities requested. However, the generation rate of source S3 has not been altered, since that source requested a lower value which is possible to attain. For the rest of the traffic types, the generation rate of all the nodes has been reduced, adjusting them to very similar values.

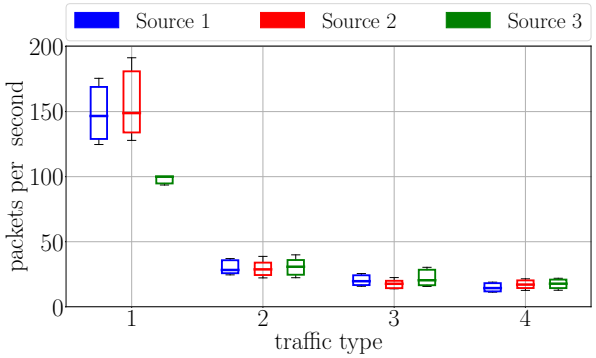


FIGURE 4.24: FDCC resulting generation rates for source nodes with unbalanced packet generation rates.

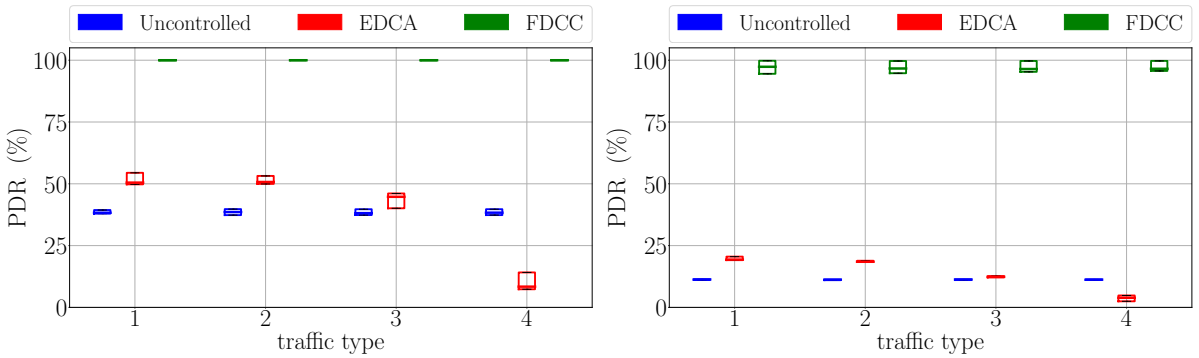
Overall, this first scenario explains the way FDCC works, and illustrates its usefulness. In the next subsection, second scenario where the number of nodes (sources and relays) is much larger is considered, and a more detailed evaluation is carried out.

4.4.2.3 Smart grid scenario

A second scenario consists of a series of network nodes arranged in a square grid, as illustrated before in Figure 3.6. One node serves as destination (data concentrator) while other nodes act as traffic sources and intermediate routers. Source nodes generate data traffic belonging to the four traffic types, and directed to the data concentrator, in the same way than in the previous scenario (see Table 4.7). Clearly, given the workload delivered by each node, the network will suffer congestion unless some form of control is made on the sources packet generation rate.

Packet delivery ratio (PDR) and network throughput

First, the packet delivery ratio for each traffic type is studied. Two different network sizes have been considered, one of them with 9 network nodes and the other with 25. Figure 4.25 shows the PDR results for the two cases. As can be seen, in the absence of FDCC solution a high percentage of packets will be lost on their way towards the data concentrator. The loss of packets is greater when the size of the network grows. Remember that all nodes try to transmit equally, and therefore as the number of nodes grows, the attempt to use wireless channels also grows, leading to a higher congestion situation. Note that packets that are lost along their way to the data concentrator have previously consumed network resources in a completely useless way. All these disadvantages can be resolved with a smart generation rate control at the source nodes, achieving a PDR near 100% as can be seen in the figure when the FDCC mechanism is applied.

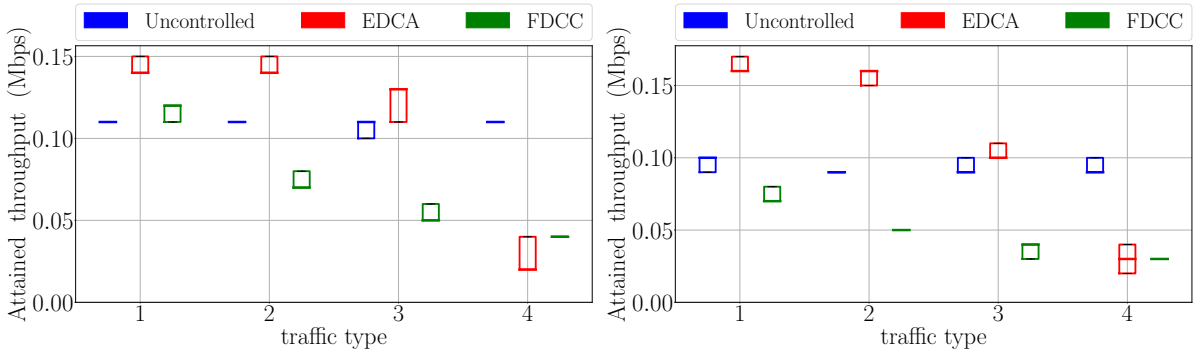


(a) Network size 9 nodes

(b) Network size 25 nodes.

FIGURE 4.25: Packet delivery ratio for each traffic type and for different network sizes.

To have a better understanding on the network performance, the attained throughput is evaluated. Figure 4.26 shows the results for the two network sizes under consideration. Again, it can be seen that the attained throughput by the different traffic type is the same when differentiation mechanisms are not applied. EDCA does provide differentiation, allocating more network resources to the most priority traffic, as does FDCC. Given the set of parameters considered for these two mechanisms in these simulations, we observe that the overall throughput attained by EDCA exceeds that obtained by FDCC. In fact, the attained throughput obtained with FDCC is even lower than the obtained without applying any mechanism. These values can of course be modified by setting less restrictive parameters for FDCC. However, and more importantly, the attained throughput with FDCC has a higher quality in two ways as will be seen below: fairness in the distribution among the different network nodes, and network transit time.



(a) Network size 9 nodes.

(b) Network size 25 nodes.

FIGURE 4.26: Attained throughput for each traffic type and for different network sizes.

Fairness in the network resources distribution

So far, the combination of Figures 4.25 and 4.26 has shown that the implementation of the proposed solution enables all applications to experience a near 100% packet delivery ratio and to differentiate their attained throughput based on their assumed criticality while maintaining a high level of the network resource utilization. As previously said, the next step is to study the fairness in the network resources distribution among the different source nodes. The point here is to observe how, in the absence of a fair mechanism, the nodes that are closer to the destination obtain a better service from the network (higher throughput) than those that are further away.

This fact is clearly seen in Figure 4.27, which shows the attained throughput obtained for each traffic type, depending on the distance from the source node to the destination (measured in number of hops). The network size considered here is 25 nodes, and as can be seen in the figure the distance from the source nodes to the destination varies between 1 and 8 hops. With both uncontrolled and EDCA modes, and for all traffic types, nodes located at a distance greater than four hops get very poor performance. In addition, as can be seen with the boxplots, the variability presented by the results is very high. However, applying the FDCC mechanism, the value of the flow rate is practically constant with regard to the number of hops, and with a very small variability in the results. Both facts are extremely important, since on the one hand the same service is offered to all users, and on the other the quality of the experience of each of them remains constant over time. This fact is especially serious for traffic types 3 and 4, which would be otherwise virtually eliminated when they originate from nodes located more than four hops away.

To complement and quantify these results, in Figure 4.28 we present the value of the Jain's index calculated on the attained throughput by traffic flows coming from different nodes. As can be seen, the value of this index is considerably higher when the FDCC mechanism is applied, with values very close to the maximum. On the other hand, when the size of the network grows, as expected the variability between nodes is even greater, so that the improvements obtained with FDCC are even more relevant.

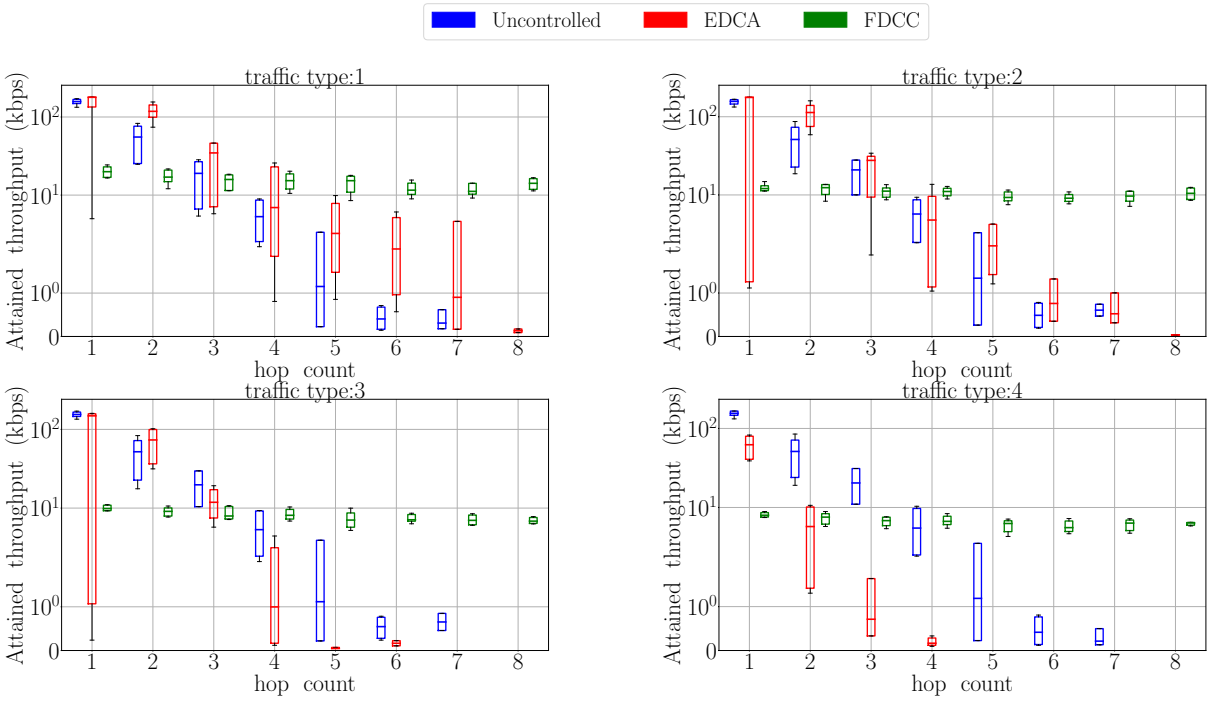
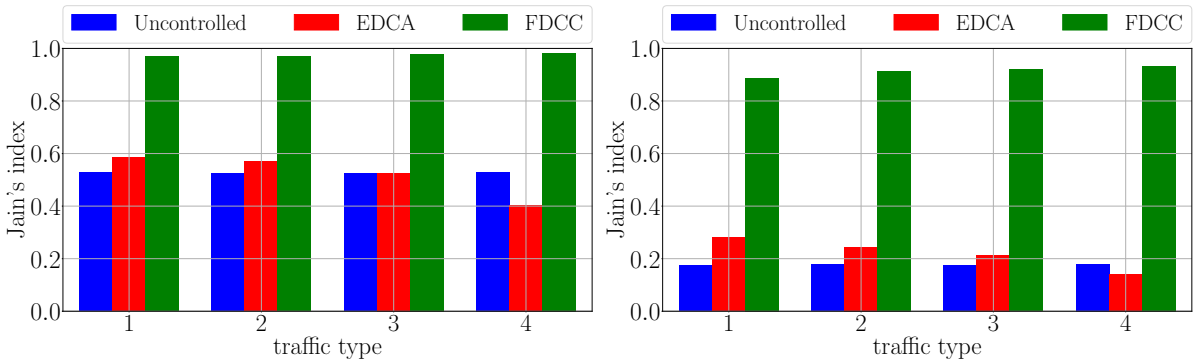


FIGURE 4.27: Throughput vs number of hops for each traffic type (Network size 25 nodes).



(a) Network size 9 nodes.

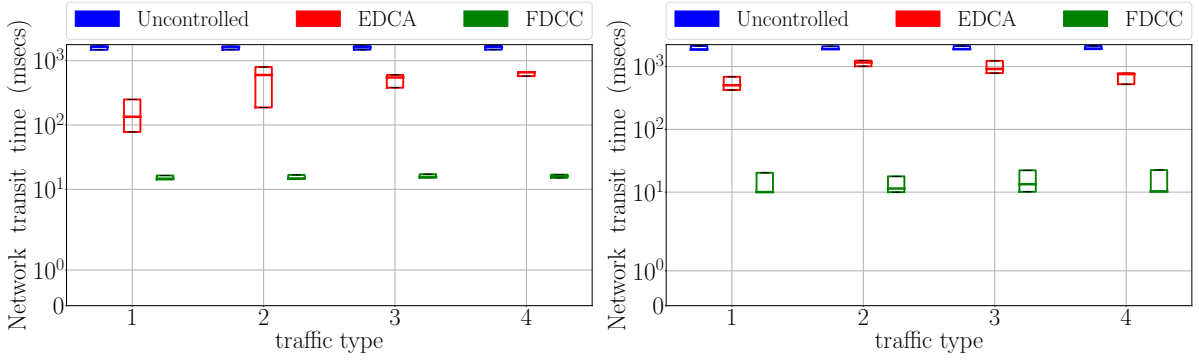
(b) Network size 25 nodes.

FIGURE 4.28: Fairness on the attained throughput for each traffic type and for different network sizes.

Network transit time

In addition to the avoidance of network congestions and buffers overflows, implementing a packet generation rate control typically leads to lower packet network transit times. This is confirmed by Figure 4.29 where both EDCA and FDCC manage to reduce this time for all traffic types, with a more significant decrease in the case of FDCC. Note also that in both uncontrolled and EDCA modes, packets from distant nodes are dropped halfway to their destination. Therefore their transit times, which would be typically large if they had arrived to their destination, are not taken into account when calculating the average value. However, with FDCC these traffics are regulated so that the obtained average values for the network transit time are still significantly lower.

Regarding the network transit time, an interesting measure is the percentage of packets of each traffic type that has been delivered to their destination in a time less than a given maximum delay. Clearly this maximum delay depends on the importance (traffic type) of each traffic flow. In our case, we use the values given in Table 4.11. We refer to the percentage of packets meeting their time constraint as the compliant factor and we represent its values in Figure 4.30. In addition to the factor for each traffic type, a global value is offered, taking into account the proportion of packets of each traffic type. As can be seen, the value of this factor is considerably better in all cases when applying FDCC, taking practically its maximum value.



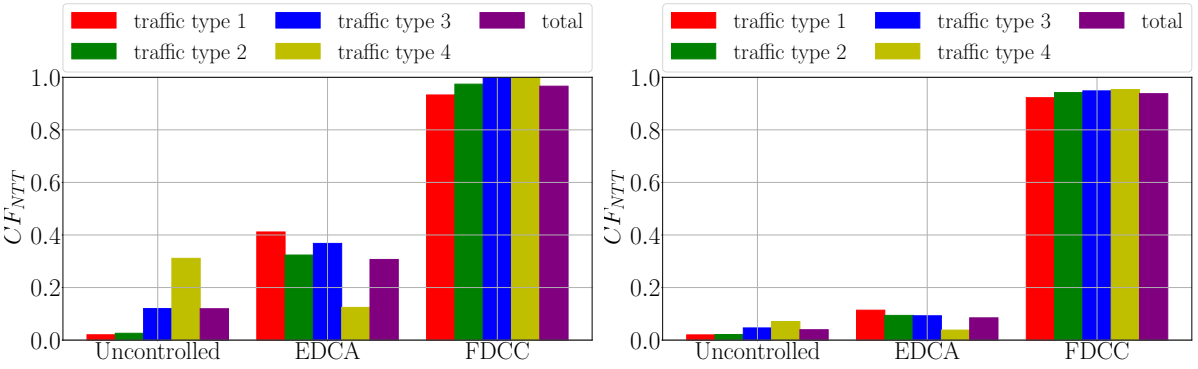
(a) Network size 9 nodes.

(b) Network size 25 nodes.

FIGURE 4.29: Network transit time for each traffic type and for different network sizes.

TABLE 4.11: Maximum allowed network transit times.

Traffic type	Network transit time
1	50 ms
2	100 ms
3	1000 ms
4	2000 ms



(a) Network size 9 nodes.

(b) Network size 25 nodes.

FIGURE 4.30: Network transit time compliant factor.

Finally, the additional control traffic introduced by the implementation of FDCC have been measured. This traffic corresponds to the packets generated by the routers to indicate to the traffic sources the generation rate that they are allowed to transmit for each category. The amount of traffic generated depends on the averaging period selected for the algorithm (T_A). But it must also be taken into account that control packets are only generated when new allowable rate values (that is, values that will modify the current generation rate) are obtained by the routers. Specifically, for all the experiments carried out, the value of this control traffic is always between 3 and 4% of the data traffic.

4.4.2.4 Downstream traffic from the control center to the NAN

As a third simulation scenario, the inclusion of downstream data traffic from the control center destined to each of the network nodes has been considered. Within the scope of the NAN, this traffic will therefore have the data concentrator as the source node. In order to demonstrate the versatility of the FDCC mechanism under various configurations, let us consider that the downstream traffic coming from the control center is of extreme importance, included in traffic type 1, so that it will not be affected by the rate control mechanism. In other words, the data concentrator has to forward all the received packets from the control center. The rest of the network nodes will generate traffic belonging to traffic types 2, 3 and 4. Note that the rate control mechanism applies to these traffic types as in the previous scenarios. Table 4.12 shows the set of parameters chosen for the new downstream traffic. Its packet generation rate depends on the number of network nodes (here 25), and has been set to 5 pkt/s per destination node. The FDCC mechanism selected parameters are similar to those presented in Table 4.9, the only difference being in the α values assigned to each traffic type. Given that the packet generation rate of NAN applications that belong to traffic type 1 is not controlled, the α values for this

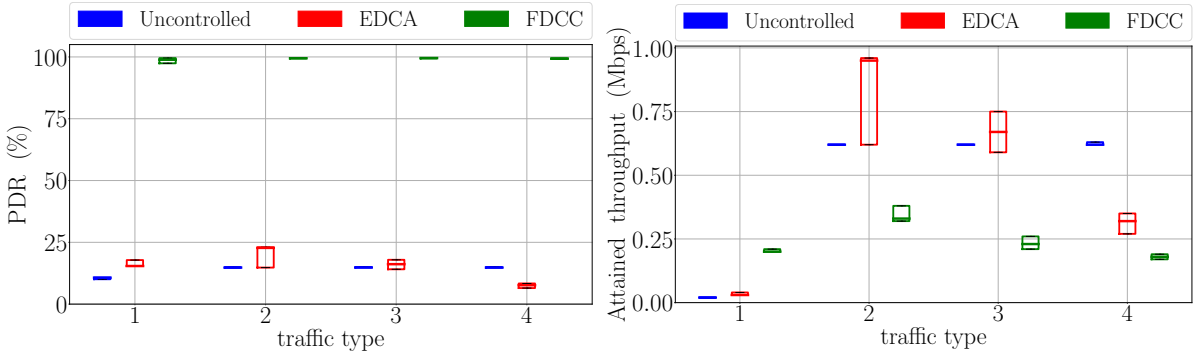
traffic type is not needed. For traffic types 2, 3 and 4, the chosen values are 0.8, 0.7 and 0.6 respectively.

TABLE 4.12: Downstream data traffic.

Traffic type	**Applications**	**Packet length average (Bytes) and PDF**	**Packet generation rate (per sec) average and PDF**
1	Demand Response, Outage Management	200 Exponential	5 Exponential

Packet delivery ratio (PDR) and network throughput

In this scenario, a network arranged in a grid with 25 nodes has been considered. The network performance is evaluated in terms of PDR, attained throughput, network throughput fairness and transit time. Figure 4.31 shows the PDR and the attained throughput. Similarly to the previous scenarios, Figure 4.31a shows how the PDR reaches almost 100% with our solution compared to the other two cases where the PDR is less than 25% for all the traffic types. Regarding the attained throughput, as can be noticed in Figure 4.31b that the value for traffic type 1 is significantly higher when FDCC is applied. In fact, the obtained value corresponds to all the throughput demanded by the data concentrator (since the generation rate of this traffic is not regulated and the PDR is 100%). For the rest of the traffic types, the FDCC mechanism tends to act likewise to the previous scenarios.



(a) PDR.

(b) Attained throughput.

FIGURE 4.31: Packet delivery ratio and attained throughput for each traffic type (Network size 25 nodes).

Fairness in the network resources distribution

The attained throughput fairness between the network nodes can be seen in Figure 4.32, where the Jain's index values are shown for each traffic type. Not surprisingly, for traffic type 1, the Jain's index has always the maximum value, since this traffic is only sent by the data concentrator. For the rest of the traffic types, the distribution of the attained throughput between the different nodes is much more uniform when FDCC is working.

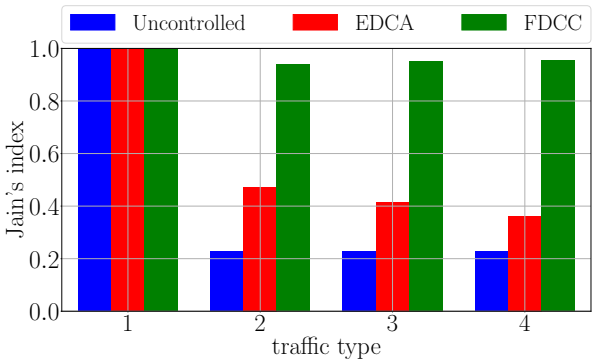
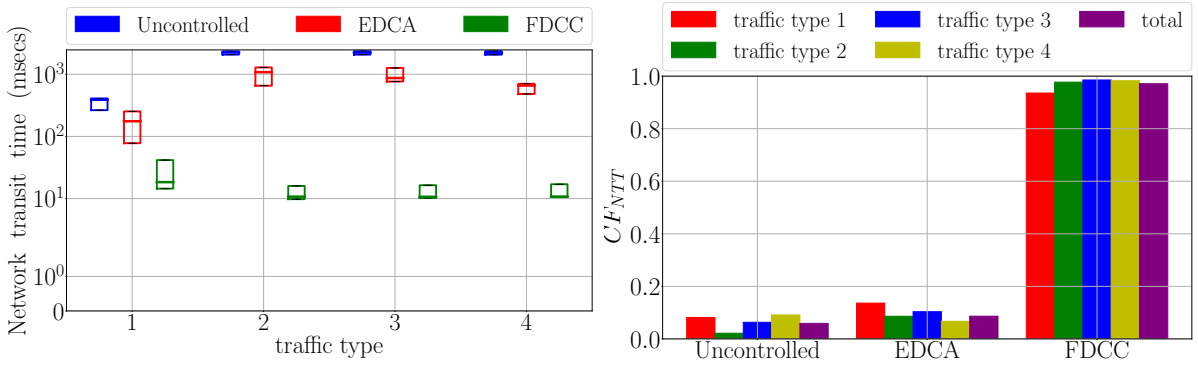


FIGURE 4.32: Fairness on the attained throughputs for each traffic type (Network size 25 nodes).

Network transit time

Finally, Figure 4.33 shows how in this scenario the improvements obtained in terms of packet network transit time are also achieved. In Figure 4.33a it can be seen how the values are lower with FDCC for all the traffic types, including the new and uncontrolled downstream traffic type 1. Regarding the compliant factor, the obtained values are depicted in 4.33b, where the maximum values selected in Table 4.11 have been taken into account. As it can be seen, the minimum requirements for the network transit times are significantly better accomplished for all the traffic types when our solution is implemented.



(a) Network transit time.

(b) Compliant factor.

FIGURE 4.33: Network transit time and compliant factor for each traffic type (Network size 25 nodes).

4.5 Conclusions and future work

In this chapter, two different congestion control solutions for wireless multi hop networks were presented. The conclusions and future line of works are described below.

EA-HWMP

Section 4.3 has evaluated the feasibility of implementing an emergency aware congestion control mechanism in the context of Smart Grid Neighborhood Area Networks. For this purpose, a Wireless Mesh Network is used as the network technology and smart meters as mesh STA devices. First, the proposed mechanisms allow identifying the different applications and grouping them into different types of traffic. Second, different congestion control functions have been configured for each traffic type. These functions are intended to give a higher transmission probability to NAN applications with higher priority in situations of network congestion. In addition, these functions are adapted depending on the emergency situation. That is, in a high emergency situation, the transmission probability is increased even more for priority traffic types by discarding applications with lower QoS requirements. Also, the proposal works together with a channel allocation technique. This mechanism separates control and data traffic into separate channels. On the one hand, a dedicated control channel is responsible for transmitting routing and emergency messages. In particular, emergency messages must be transmitted to the whole network in the shortest time possible, and therefore a dedicated channel has been used to avoid unnecessary delays that may be caused by network load on the data channels. On the other hand, the rest of available channels are used to transmit the different NAN applications. The objective of EA-HWMP proposal is to use the least amount of data channels, while keeping the desired QoS for all traffic types. For this, one more channel will be used to transmit the data if and only if the current channel is busy ($\rho_{ch} > \rho_{th}$). Finally, if all the channels are busy, the packets are discarded.

In order to evaluate the EA-HWMP proposal, two network scenarios have been considered (congestion control and emergency). Besides, the number of nodes present in the grid is increased (from 9 to 36 nodes), which gives rise to a greater contention in the access to the shared medium. First, the congestion control scenario is evaluated with three different network loads (normal, medium and high), and without using the emergency aware mechanism. For this scenario, the number of the data flows for traffic type 1 is increased during the medium network load, while the packet generation rate for the other traffic types are remained the same for all the simulation. From the obtained results, it can be seen that in situations of network congestion, our proposal mechanism overcomes the basic protocol in terms of packet delivery ratio, throughput and transit time for priority traffic types and for all network sizes. Second, with the emergency scenario, the network performance has been evaluated when the channel allocation and traffic differentiation techniques, congestion control mechanism and the emergency system are implemented together. In this scenario, NAN applications were configured to generate the same amount of data for all traffic types, and also four emergency sub-scenarios are considered (normal, medium, high and a combination of the previous three). These emergency situations were propagated through the whole network by using broadcast messages. It can be seen from the results that emergency aware mechanism benefits the traffic with higher QoS requirements in situations of high emergency. It can be noticed clearly in the combined emergency sub-scenario how the PDR and the delivered throughput are modified according to the emergency situation configured.

FDCC

Section 4.4 has presented a conceptually simple congestion control mechanism for an IEEE 802.11-based NAN that carries traffic from several types with different QoS requirements. As it was mentioned before, NANs are typically multi-hop networks so that handling priorities only

at a link-level would undoubtedly lead to favoring nodes close to the data concentrator (NAN gateway). FDCC relies on three distributed algorithms executed on the Smart Meters that work together to adequately regulate the traffic sources with regard to their traffic type and hence avoiding network overload.

The proposed solution easily accounts for different traffic types (with different QoS) while ensuring a fair sharing of the network resources between them and, importantly, an access to the network resources even for sources far from the data concentrator. The results has been compared to the EDCA mechanism that works only at the link-level. FDCC has shown that is efficient in terms of packet delivery ratio and transit time. While its performance at network throughput may appear a bit lower than EDCA, this is only the result of a better fairness in the sharing of network resources. As a future work, an emergency system will be implemented. For this purpose, a dynamic allocation of α will be proposed as well as an emergency signaling.

Going further in the search for better performance of congestion control mechanisms in multi-hop wireless networks, future line of works are the following. The applicability of the algorithms proposed in this chapter in other communications network environments will be studied, such as in crowded networks with mobile nodes. These networks can be built for example through the intercommunication of cell phones of people who roam their cities. Here, all the network nodes (or a selected group among them) could periodically broadcast the allowable maximum generation rate as a function of the measured channel utilization factor and the number of 1-hop and 2-hops neighbors.

In the next chapter, a different machine learning algorithm for congestion control on NANs will be proposed, presented and evaluated.

Chapter 5

Machine Learning Congestion Control mechanism

We are living in an era where data comes in abundance, and many research areas are focused on making sense of this data. In this context, machine learning has evolved as a field of the Artificial Intelligence (AI) where different algorithms are capable of generating knowledge through data, or in other words, make predictions. In machine learning, datasets are historical data that serve as the the basis for training the predicting model. The objective is that a model make decisions from the data provided by dataset. Keep in mind that these historical data can comes in different ways: structured or unstructured data. Therefore, some preprocessing techniques are necessary before training the model.

In this chapter a different congestion control mechanism for wireless multi-hop networks which is based on self-learning algorithms is presented. Specifically, a decision tree classifier is implemented that, in different network congestion situations, can predict if a packet must be transmitted or not according to its priority. This proposal is evaluated again in the context of a NAN topology made up of smart meter devices where the selected technology is the wireless mesh networks.

Throughout this chapter, it will be emphasized that building the dataset correctly plays a fundamental role in the precision and accuracy of the model. That is, the quality of the data and the amount of useful information are prevalent factors on how well a machine learning model can learn. Therefore, how to create a good dataset for multi-hop wireless networks will deeply covered.

5.1 Introduction

Machine learning has been exploited widely in many applications due to the large amount of data available today, and it has been proven to be a system capable of solving problems from the simplest to the most complex. Nowadays, the computing power has increased in terms of number of central processing units (CPUs), graphics processing units (GPUs) and tensor processing units (TPUs), so it is possible to use advanced learning models such as representation learning, distributed and parallel learning, deep learning, and so on [92]. However, the basis for a machine learning model is the dataset, and depending on the dataset, a machine learning

technique will perform better than other. For this reason, it has been considered of importance to describe the most important parts of a dataset by using a simple example.

Datasets are groups of data that can represent anything. Figure 5.1 shows the well-known iris dataset[1] which it is composed for the measurements (samples or observations) of 150 iris flowers. Each sample contains the features (attributes, measurements, dimensions) and the labeled class (target). Besides, features are the elements used to take decisions when training the model. For instance, the flower measurements will be the inputs of the model (features), and the flower specie will be the expected outcome (class).

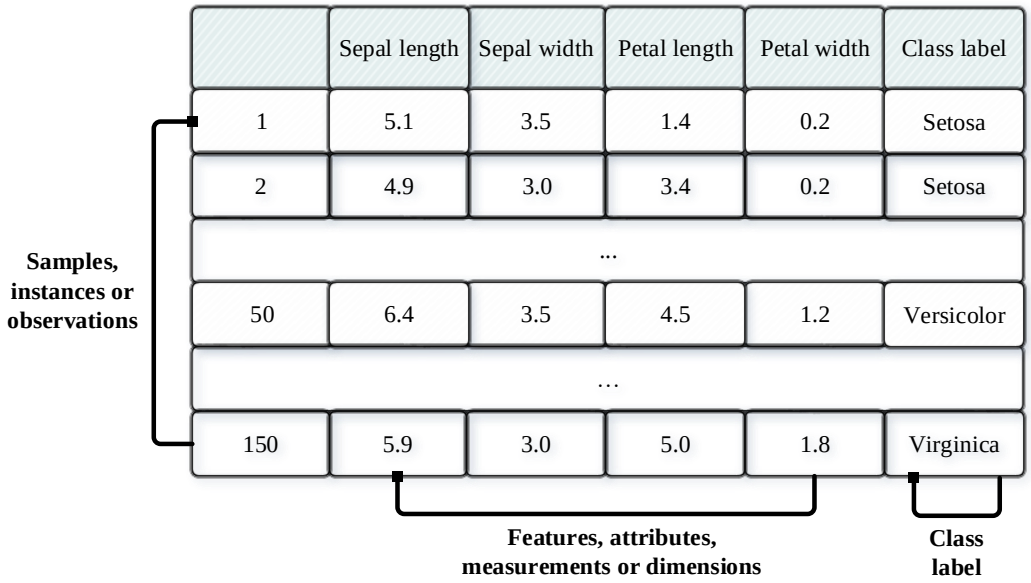
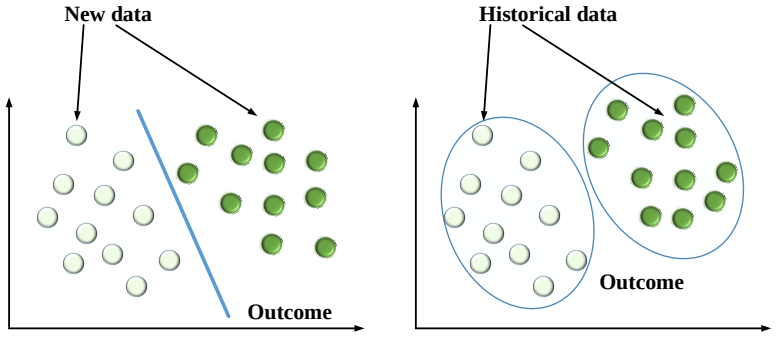


FIGURE 5.1: Iris dataset.

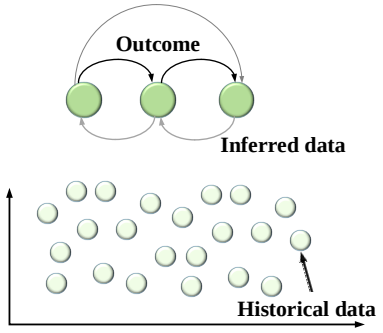
There are different machine learning techniques, but they are mostly classified into three types: supervised, unsupervised, and reinforcement learning (see Figure 5.2). Each of one is targeted to solve an specific problem, and they are briefly described below:



(a) Classification. (b) Clustering.

FIGURE 5.2: Examples of supervised, unsupervised, and reinforcement learning algorithms.

[1] https://archive.ics.uci.edu/ml/datasets/Iris



(c) Rule extraction.

FIGURE 5.2: Examples of supervised, unsupervised, and reinforcement learning algorithms (cont.).

- **Supervised learning:** Involves a set of algorithms that predict future outcomes based on past observations. In supervised learning, input and output pairs are used to train the model. Basically, supervised learning is used for those cases where the inputs and outputs are known in advance (labeled data). The objective is to generate an approximated model that predicts outputs to unseen data. Supervised learning algorithms are divided into two groups: classification and regression problems [93]. For instance, the classification of flowers presented above is an example of supervised learning. That is, depending on the flower measurements, a suitable algorithm will predict the specie. Besides, these models take a direct feedback to verify if the predicted output is correct or not.

- **Unsupervised learning:** Involves a set of algorithms that infer patterns or structure from unlabeled input data. Unsupervised learning are aimed to solve clustering and association problems [94]. In order to understand this statement, consider again the example of the iris data set. When training with an unsupervised algorithm, only the input dataset will be provided, and with the help of an appropriate algorithm, the model will find the patterns and separate the flowers into different groups. Besides, these algorithms do not need any feedback.

- **Reinforcement learning:** They are a set of algorithms which takes decisions based on an agent that interacts with the environment, where their actions are rewarded or penalized [94, 95].

In this chapter, a supervised learning algorithm for congestion control in multi-hop wireless network will be proposed, implemented and evaluated. The objective is to classify if a packet will be correctly received or not based on the current network load. The scenario of evaluation is again the Smart Grid Neighborhood Area Networks where the selected technology is the wireless mesh network made up of smart meter devices. For this purpose, a dataset for NAN will be generated from scratch. As in with previous proposals, the solution will provide also traffic differentiation. To this end, an approximation to generate a meaningful dataset that covers both different network loads and traffic differentiation will be also considered. Besides, the generated model will be evaluated with machine learning performance metrics as well as with the ns-3 simulator.

5.2 Proposed solution

Figure 5.3 presents the road-map used for building our Machine Learning Congestion Control (MLCC) mechanism. The proposed solution is aimed to the wireless multi-hop network and specifically for the Smart Grid Neighborhood Area Networks. In order to generate the machine learning model the following tasks has to be carried out: data collection, data processing, validation and evaluation of the model.

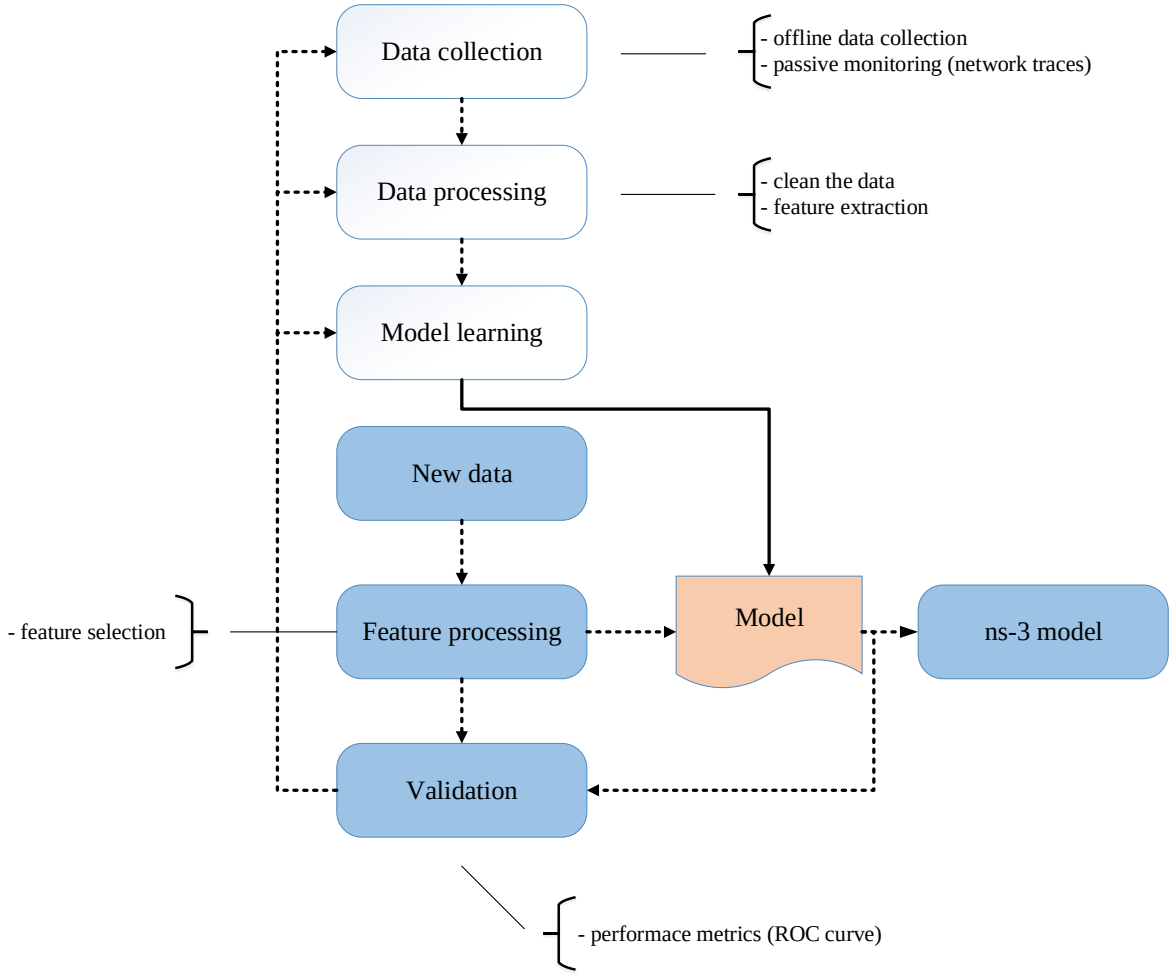


FIGURE 5.3: Machine learning system [1].

First, data collection is mandatory since there is not an available dataset for NAN that contemplates different network loads and traffic differentiation. To this end, different network parameters that describes the network state will be collected. For our purposes, the channel utilization factor and buffer occupancy are good descriptors for network congestion. In data collection stage, several simulation runs will be performed taking into account different network loads generated for different NAN applications.

Secondly, the collected data has to be processed before training. That is, the features and classes are extracted, in other words, inputs and possible outputs of the training dataset has to be defined. The proposed solution belongs to the supervised machine learning models, and specifically is a binary classifier because the possible outcomes are predefined. That is, the model

classifies whether the packet will be transmitted correctly or not to the destination based on the current network state. For this purpose, the decision tree classifier will be used. This classifier is easy to understand and implement. Besides, its prediction power will be evaluated to unseen data with machine learning performance metrics such as the Receiver Operating Characteristic (ROC) curve.

Finally, the generated model will be evaluated in the ns-3 simulator with different Smart Grid Neighborhood Area Network scenarios.

In the next subsections, the data collection and processing, validation and implementation in the ns-3 simulator of the machine learning model wil be deeply covered.

5.2.1 Data collection

The basics for the machine learning are the data. Besides, data collection is mandatory when there is not available dataset. There are many datasets available online, and each one contains information about anything such as demographic information, wines classification, climate behaviour, and so. Therefore, the dataset is strongly related to the targeted problem to solve.

One of the objectives of this dissertation is to propose a congestion control mechanism for wireless multi hop networks. Besides, it is necessary to have meaningful data that represents different network loads. Note that there is not an available dataset for our targeted problem. Therefore, several networks parameters will be collected. Before detailing the training process, the data collection scenario will be explained.

Figure 5.4 shows a Smart Grid NAN scenario consisting of eight smart meters (nodes 1 to 8) and one data concentrator (node 9). In this data collection scenario, it has been considered only the upstream traffic from the smart meters towards the data concentrator. Data collection is a very important step to create the learning model, and defining which parameters will be the best descriptors of the network state is essential. It has been concluded from previous chapters that the channel utilization factor (ρ) is a good parameter to describe how congested is a node. Besides, the buffer occupancy (q) or number of queued packets has been also considered as a representative network congestion feature to generate the training dataset. Note that these two parameters are strongly related.

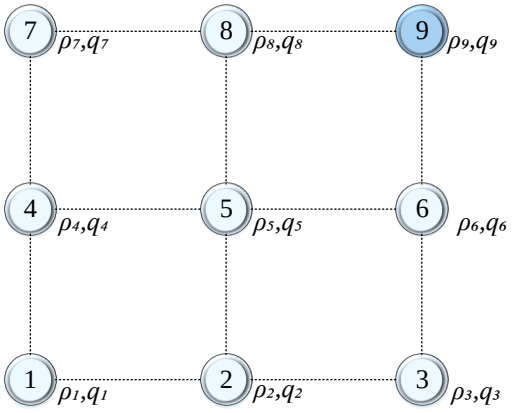


FIGURE 5.4: Data collection scenario.

The data collection process can be carried out in two different ways: offline or online. In our context, offline data collection will be performed, so a large amount of historical data about the network state will be collected by monitoring different parameters. For instance, the channel utilization factor, buffer occupancy and arrival packet times among other parameters are collected. Besides, the offline data collection can be done through an active or passive monitoring. For our purposes, passive monitoring will be performed in order not to introduce additional control traffic such as probe packets [95].

Passive monitoring is done through ns-3 network traces with the objective of collecting all the traffic transmitted over the NAN and the network events. These monitoring tools capture data at IP, MAC and physical layer. The network traces used for the data collection are presented in Table 5.1. With these network traces, the most relevant events related to the packets are collected (packet transmission and reception hop by hop) as well as the network congestion state parameters (channel utilization and buffer occupancy).

TABLE 5.1: Network traces used for passive monitoring [2]

	Network trace	**Description**
Physical layer	State	The state of the PHY layer
MAC layer	MacRx	A packet has been received by the current device, has been passed up from the physical layer and is being forwarded up the local protocol stack.
	MacTx	A packet has been received from higher layers and is being processed in preparation for queueing for transmission.
IP layer	IpTx	Send ipv4 packet to outgoing interface.
	IpRX	Receive ipv4 packet from incoming interface.
Queue	Enqueue	A packet arrived at the MAC for transmission.
	Dequeue	A packet was passed down to the PHY from the MAC.

Several simulation runs were performed to built the training dataset. The main simulation parameters used to collect the data are presented in Table 5.2. Furthermore, a variable size (specifically, an exponential truncated distribution) has been chosen for the data packets. With respect to the inter arrival time average, the same criteria has been followed, and so packets are generated with different and random time between each other. In order to have the full range of values of the channel utilization factor and buffer occupation, different inter arrival time average values have been selected for each random seed and run number.

TABLE 5.2: Main simulation parameters used for data collection.

Description	Value
Network simulator	ns-3.28
Distance between nodes	80 m
Simulation time	170 s
Tranport layer	UDP
Random number generator	MRG32k3a
Random seed	[1 . . . 14]
Run number	[1 . . . 21]
Routing protocol	HWMP
Routing metric	ALM
Maximum queue size	100 packets
Inter arrival time average and PDF	from 0.01s to 0.3 s (Exponential)
Packet length average and PDF	200 Bytes (Exponential)

Figure 5.5 shows an example about how the data collection has been done. For this, one packet that belongs to traffic class 1 is transmitted from smart meter 1 towards the data concentrator. As can be seen, when the packet transverse the network hop by hop, the current channel utilization factor and the buffer occupancy are store in the repository. Besides, many more events are stored such as the arrival time, type of traffic and a packet identifier among others.

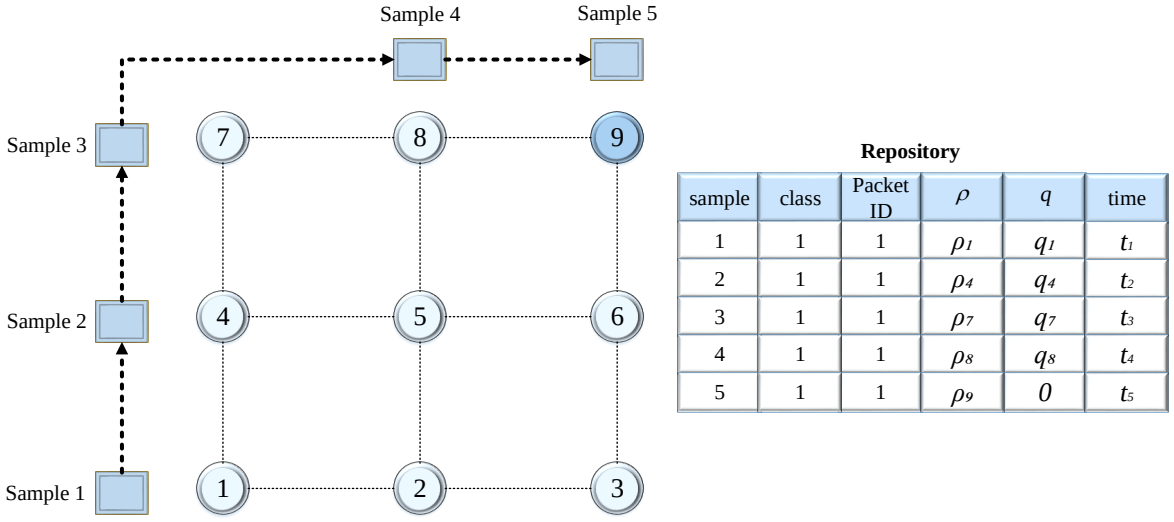


FIGURE 5.5: Passive monitoring by using network traces..

5.2.2 Data processing and feature extraction

The collected data is stored in repositories, and they are unstructured or disorganized. Therefore, the collected data has to be processed in order to organize and define the features and labeled classes that will be the inputs and outputs respectively to train the machine learning model. This process is often called as feature extraction. For this purpose, the pandas tool [96–98] has been chosen to process, create and organize the variables from the collected data in a structured way.

First, the data is processed in order to have one meaningful sample per packet. That is, each sample will contain the network parameters perceived by the packet when it transverses the whole network hop by hop. The variables used to represent each sample are presented in Table 5.3. Second, the features and classes are extracted from these variables. Figure 5.6 presents the resulting training dataset where it is shown the inputs and the outputs for the machine learning algorithm. Each training sample represents the parameters for a unique packet transmitted over the network. The features are the channel utilization factor and the buffer occupancy per node, while, the successful packet reception is the labeled class. Note that the transit time has been used to decide if a packet has been successfully received. Therefore, a packet is considered as successfully received if it arrived at the destination based on its QoS requirements (reception and delay).

TABLE 5.3: Variables used to represent the unstructured data.

Variable	**Description**	**Value**
packetID	Packet identifier	unique integer value
source	Node that generates the packet	
destination	Destination node of the packet	
class	Traffic type	[1, 2, 3, 4]
ρ_j	Channel utilization factor of node j	[0 to 1]
q_j	Number of queued packets of node j	[0 to 100]
transit	Network packet transit time	
rec	Boolean value that represents if the packet was received or not in the destination	[0, 1]

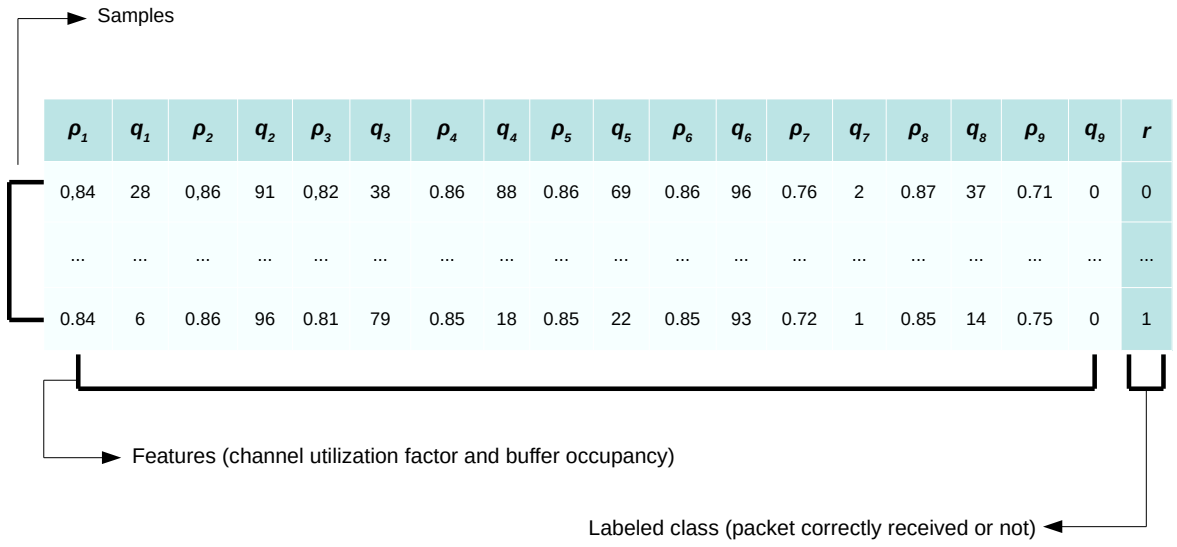


FIGURE 5.6: Features and labels for the training and testing dataset.

With these guidelines on how to build a representative dataset for our targeted problem, several data files have been generated from scratch. To this end, the scenario presented previously in Figure 5.4 has been considered. In order to generate the first dataset, the smart meters

transmit just one traffic type to the data concentrator. Other experiments will later consider the inclusion of different traffic classes. Besides, different traffic patterns have been considered to generate the dataset with the aim of collecting a large amount historical data with different values of features. As a resulting training dataset, Figure 5.7 shows how the ρ values vary from 0 to 0.8 for all nodes. It must be remembered from previous chapters that this value never reaches the value of 1.

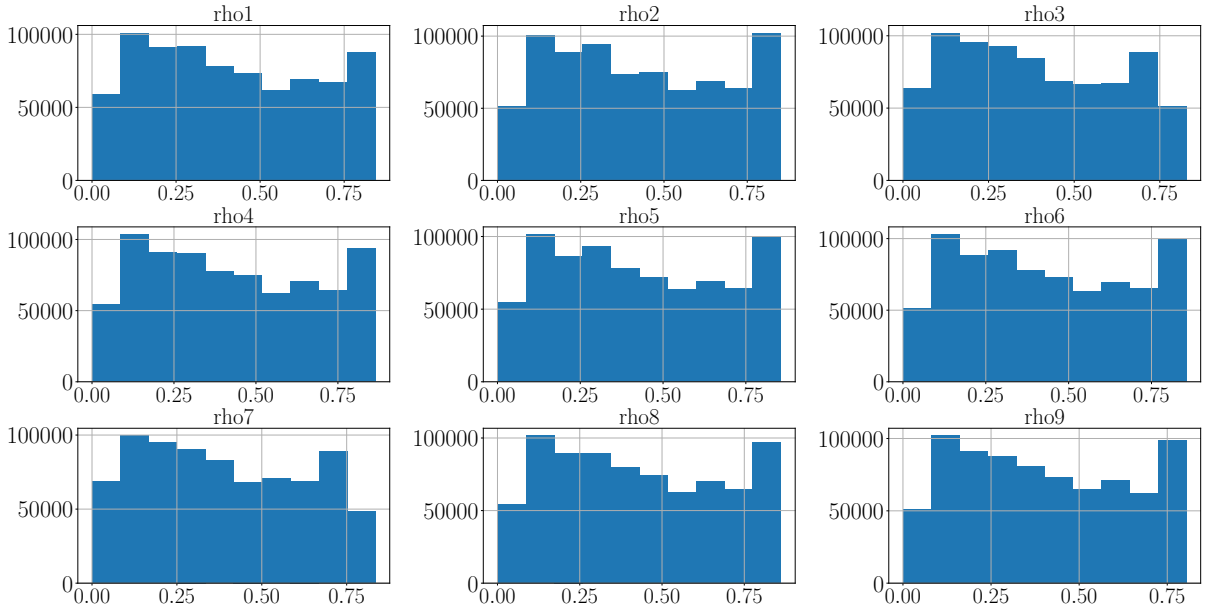


FIGURE 5.7: Channel utilization factor histogram for each node.

5.2.3 Model learning

Decision trees belong to the set of supervised learning algorithms, and they are used to solve classification problems. As the proposed congestion control mechanism is aimed to solve a binary classification problem, this type of classifier represents a very good option to take into account, with a low computational load, which is very interesting if the network nodes are based on low-cost hardware architectures. With this classifier, the prediction is approximated with a set of if-then-else decision rules [3].

Decision trees have different parameters that can be adjusted and give rise to different performances. Among those parameters, one of the most important is the depth of the tree. In general, a greater depth suppose a greater accuracy (defined as the proportion of the number of correctly predicted samples with respect to the total number of samples), although it also results in a greater number of if-then-else operations. In this work, a good trade-off has been found between the required number of operations and the obtained accuracy by selecting a maximum value for the depth of the decision trees equal to 10. Other parameters used in the classifier evaluation are shown in Table 5.4. Using these values, the resulting Receiver Operating Characteristic (ROC) curve has been obtained as shown in Figure 5.8. The ROC curve is one of the most well-known performance metrics for machine learning classifiers [99]. Basically, it is a two-dimensional graph which plots the *true positive rate* (TPR) against the *false positive rate* (FPR) for different values of the discrimination threshold:

- The TPR (also called *sensitivity*) represents the proportion of positive samples that were correctly classified as positive. It can be calculated as follows:

$$TPR = \frac{TP}{TP + FN} \tag{5.1}$$

where TP (*true positive*) is the number of positive samples correctly classified as positive, and FN (*false negative*) is the number of positive samples incorrectly classified as positive.

- The FPR (also called *fall-out*) represents the proportion of negative samples that were incorrectly classified as positive. It can be calculated as follows:

$$FPR = \frac{FP}{FP + TN} \tag{5.2}$$

where FP (*false positive*) is the number of negative samples incorrectly classified as positive, and TN (*true negative*) is the number of negative samples correctly classified as negative.

TABLE 5.4: Decision tree classifier parameters [3].

Parameters	Description	Values
Criterion	The function to measure the quality of a split.	entropy
Depth	The maximum depth of the tree.	10
Training data set size	Dataset percentage that will be used to train the classifier	0.7
Testing data set size	Dataset percentage that will be used to validate the classifier	0.3
Stratify	The training and testing dataset have the same proportions of class labels as the input dataset.	Yes

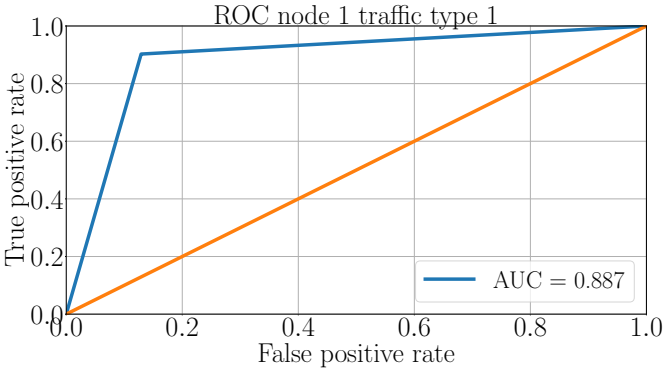


FIGURE 5.8: Decision Tree Receiver Operating Characteristic (ROC)

Therefore, classifiers that produce curves closer to the top-left corner obtain a better performance. Obviously, the best classification is achieved in the point (0,1). A uniform random prediction will provide points over the diagonal line. Points under that line represents a prediction worse than a random classification. The Area Under the Curve (AUC) is also used as a

performance value, with a maximum value equal to 1. This value is also shown in the Figure. As can be seen, the decision tree classifier trained with our proposed dataset for one traffic category provides a ROC curve with a high level of prediction performance.

For the previous performance evaluation, the simplest case has been considered in which all the network nodes in use the same classifier each time they have to make the decision of whether or not to transmit a data packet. In order to obtain a performance improvement, the dataset can be granularized, that is, divide the whole dataset into N (number of nodes) different subsets $DSS_i, i = 1..N$. In each subset, only the samples corresponding to the packet transmission of node i will be included. This way, a different decision tree will be adjusted for every node.

Figure 5.9 shows the ROC curve for the node number 1 when the dataset is split by nodes. Comparing Figures 5.8 and 5.9, the second model exhibits a better prediction power. Besides, the AUC value is higher because the second plot is nearer to the top-left corner. The cost to pay for this performance improvement is a higher computational load and an increase in the network control traffic.

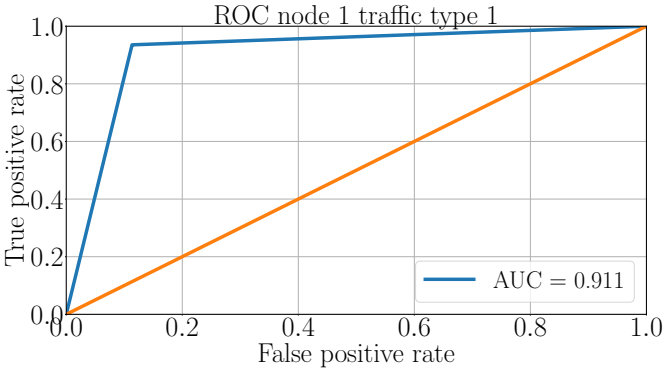


FIGURE 5.9: Decision Tree ROC with split dataset

In the following subsection, these two possibilities will be evaluated in a networking scenario with the ns-3 simulator, but first, the machine learning model implementation will be covered.

5.2.4 Machine learning implementation in the ns-3 simulator

As it was explained above, the self learning algorithms were training and validated with performance metrics such as ROC curves. However, the purpose of this dissertation is to propose a congestion control mechanism for a real data network. Therefore, the performance model will be evaluated in a more realistic environment through network simulations. For this purpose, the learning models will be exported to c++ code. To this end, the DecisionTreeToCpp tool [100] was modified in order to export automatically the different models to ns-3 simulator syntax. Note that there is a specific model per node, and thus, each exported model will have its own decision rules.

Figure 5.10 shows the machine learning framework implemented for multi hop wireless network. This figure shows how the data collection is done through network traces where different simulation runs (serially or parallel) are conducted. Besides, the collect data is process with analysis scripts. The objective is to clean the collected data and organized them in a structured way. Then, the features and classes are extracted. Next, the training stage is carried out where the tuning of the parameters and the validation are done. Finally, the model are exported to c++ code.

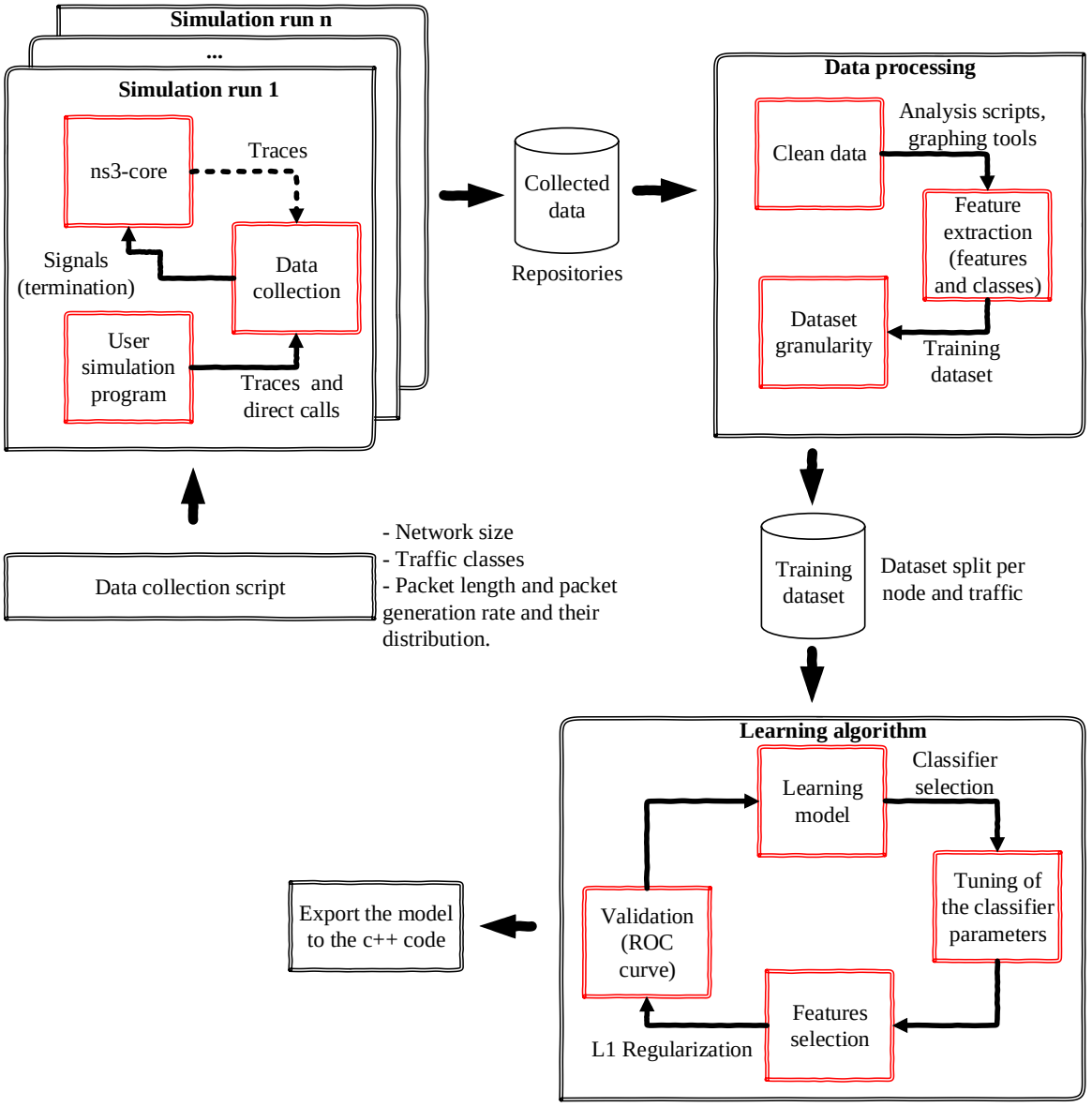


FIGURE 5.10: Machine learning framework.

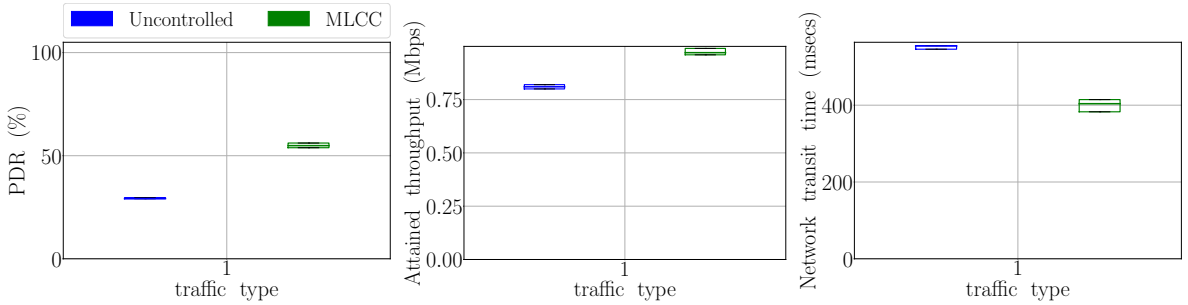
Validation of the classsifier with ns-3 simulator

The first two experiments carried out are related to the evaluation of the training dataset. On the one hand, the complete dataset will be used to generate the model. Remember that in this case all smart meter will have the same decision rules. On the other hand, the training dataset will be divided by node. That is, each node will have its own decision rules. For this two experiments, a Smart Grid NAN scenario made up of eight smart meters and one data concentrator (as already seen in Figure 5.4) has been considered. Remember also that in this first evaluation, only one traffic class is transmitted upstream from the smart meters towards data concentrator. The packet length and the packet generation rate distributions have been selected with the same values used for the generation of the training data set (Table 5.2).

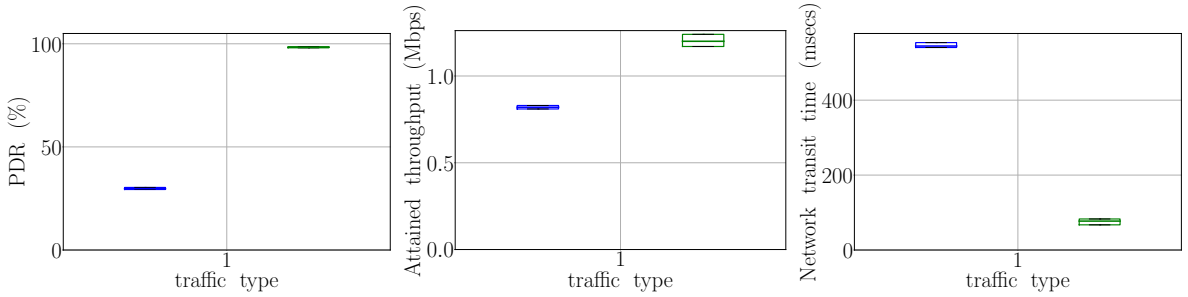
Figure 5.11 presents the results of these two experiments in terms of packet delivery ratio, network throughput and transit time. As it can been seen from the plots, the network performance

is better than the basic HWMP when the MLCC mechanism is applied.. Besides, the prediction power is even greater when the training data set is split by node. For instance, in the absence of a congestion control mechanism the PDR only peaks 26%. However, when the machine learning model is applied, the PDR reaches 50% and this value increases to 98% when the data set is split by node. Similar results are presented with network throughput where the machine learning classifier exhibits better results. That is, more information is correctly transmitted to the data concentrator. Finally, the network transit time is decreased more than a half with the proposed solution.

From these preliminary results, it is possible to conclude first that the proposed mechanism will effectively improve the network performance. Second, dividing the data set by nodes represents an even greater improvement, although it also implies a greater operational complexity in the network.



(a) Dataset complete.



(b) Dataset per Node.

FIGURE 5.11: Packet delivery ratio, network attained throughput and transit time. Network size 9 nodes.

5.2.5 Features selection

Overfitting is a common problem in machine learning, where a model performs well on training data but does not generalize well to unseen data. If a model suffers from overfitting, it also has a high variance, which can be caused by having too many parameters that lead to a model that is too complex given the underlying data [1]. The previous evaluation experiments have been carried out using a high dimensional training dataset with many features that can be irrelevant in some cases.

In order to reduce the dataset dimension as well as the model complexity, regularization techniques will be used. Regularization allows to reduce generalization error (prediction error to unseen data) by introducing a penalty for complexity. L1 regularization is typically used for feature selection which is important for our case of analysis because a large number of features

are available. With this technique, only the most relevant features for training the data will be used, and thus the model complexity is reduced (less number of rules).

With regularization, a ranking has been created where the different features of the dataset are sorted by their relative importance. Figures 5.12 and 5.13 presents the features importance at each each node, and for the previous two experiments (dataset complete and dataset split by node). Note that the feature importance plots are normalized. Besides, Figure 5.13 shows how when the dataset is split by node, in most of the cases, the most relevant features at each node are their own features. However, when the whole dataset is used by all the network nodes (Figure 5.12) the most important feature is related to the data concentrator.

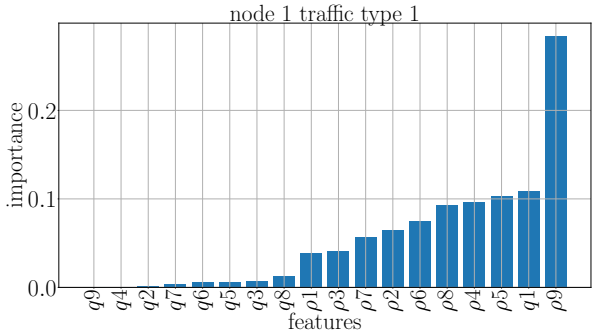


FIGURE 5.12: Features importance (dataset complete).

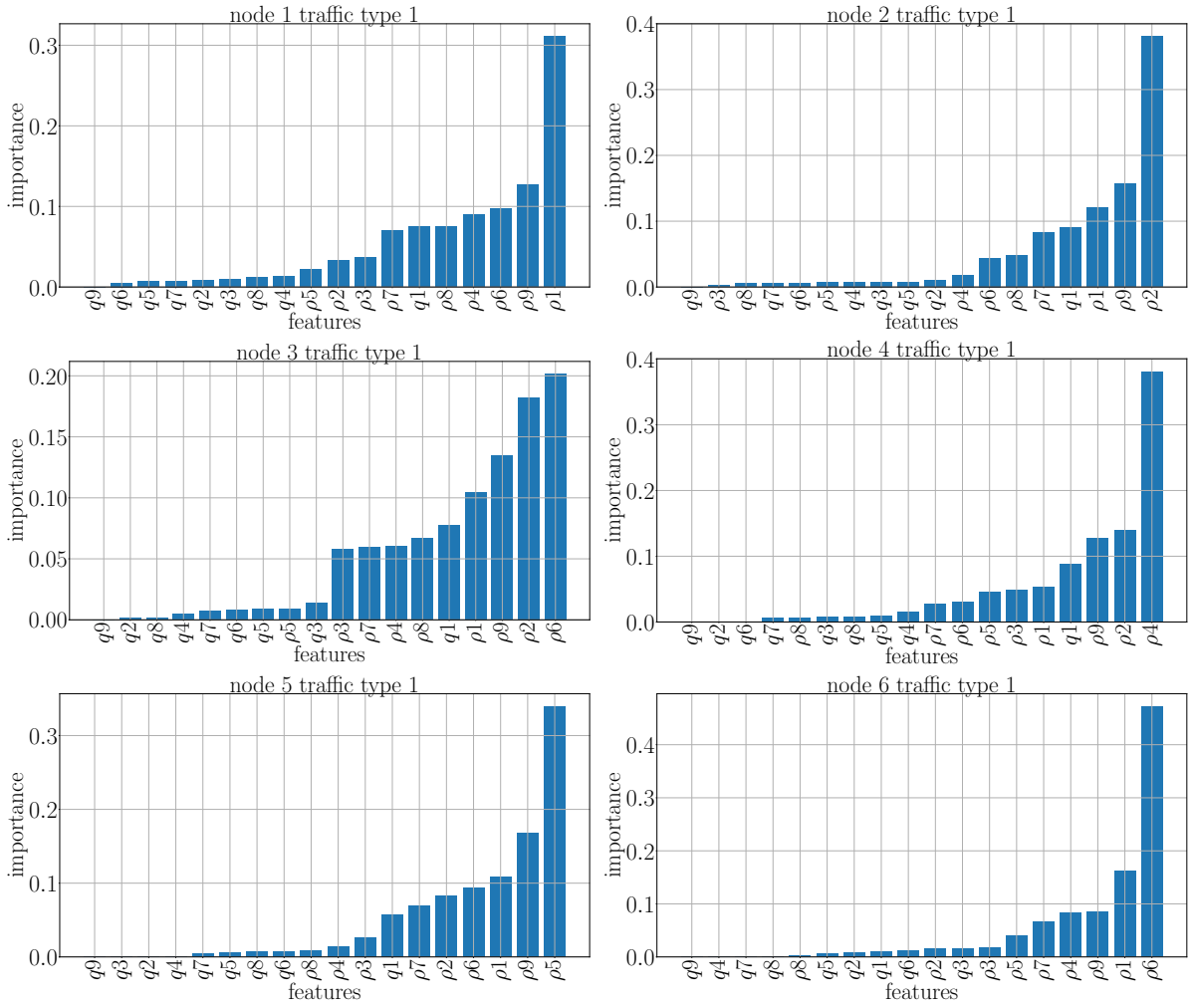


FIGURE 5.13: Features importance (dataset split by node).

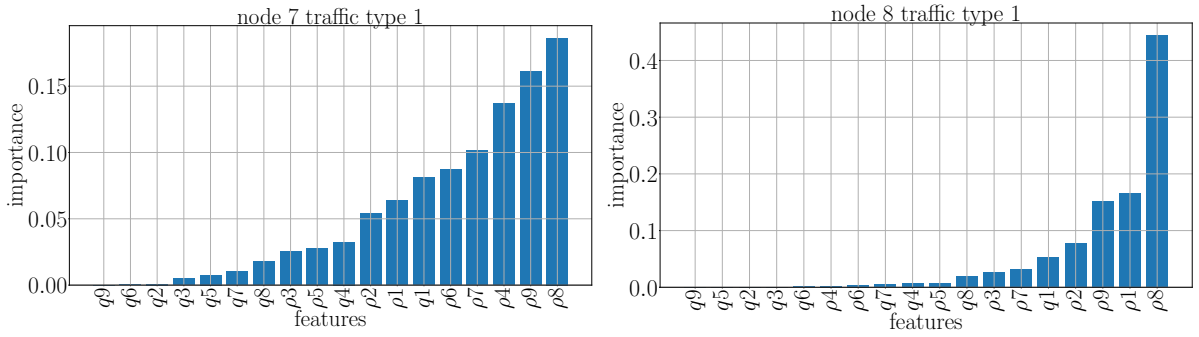
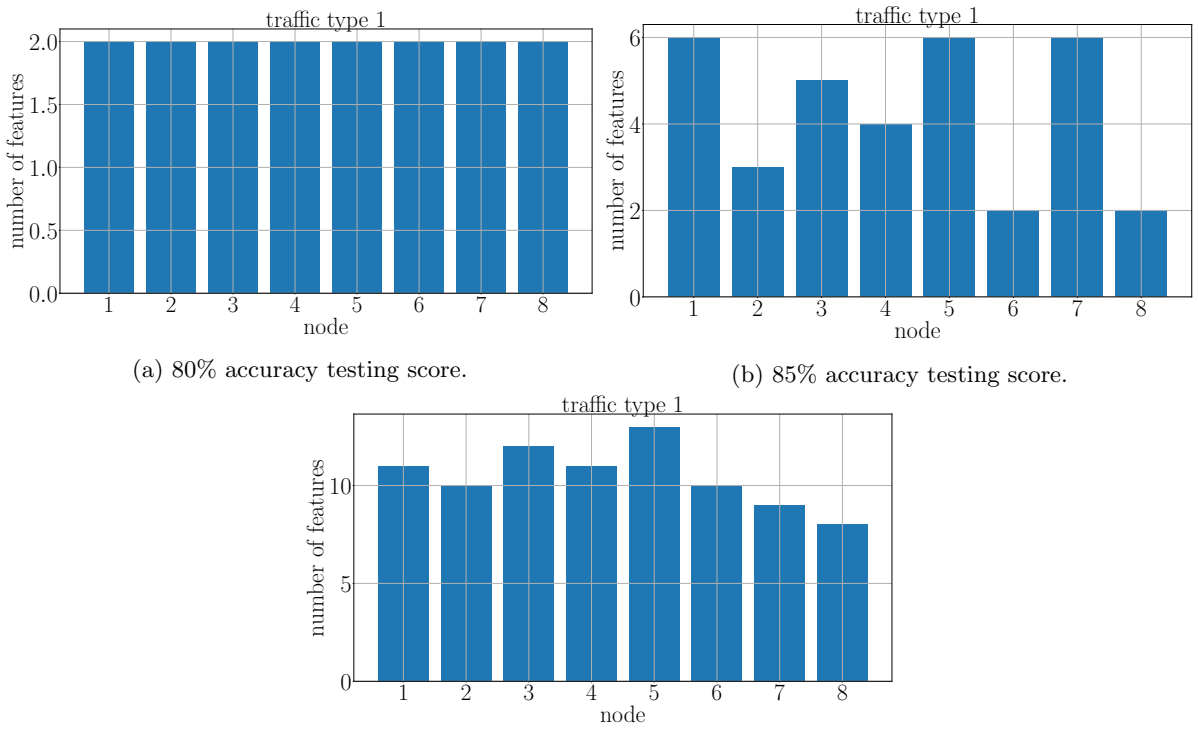


FIGURE 5.13: Features importance (dataset split by node) (cont.).

On the other hand, Figure 5.14 shows the number of features needed to reach a certain accuracy score (80, 85 and 90%). As expected, that the higher is the target accuracy score, the more features will be needed.



(a) 80% accuracy testing score.

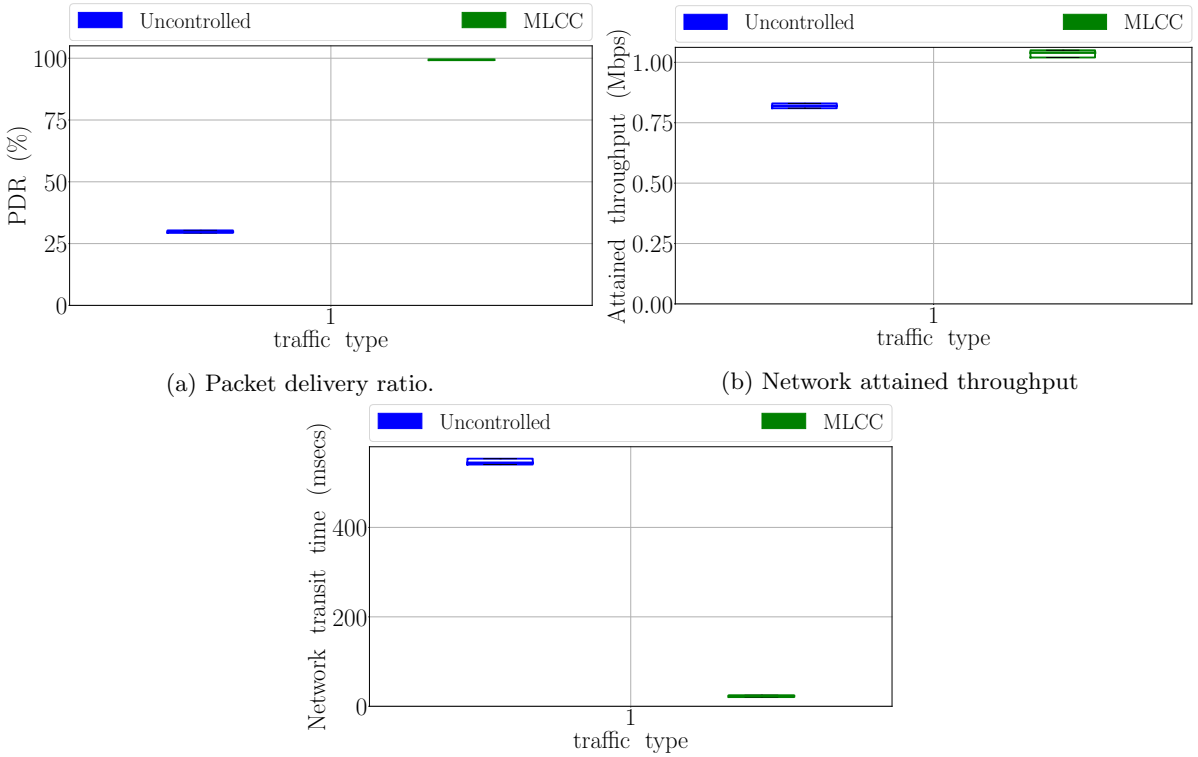
(b) 85% accuracy testing score.

(c) 90% accuracy testing score.

FIGURE 5.14: Number of features needed to reach the target testing score, Network size 9 nodes.

In order to illustrate the benefits obtained by applying feature reduction, a new set of simulations has been carried out. In this case, the models (per node) are generated with the most important characteristics to obtain an accuracy score of 85%. At the beginning of the training phase, the model is trained with the most relevant feature, and if the model does not meet the target testing score, the second most important feature will be added for the new training and so on. Moreover, a maximum number features of 6 was set up as a trade-off between a good accuracy and complexity reduction of the model.

Figure 5.15 shows the results in terms of packet delivery, throughput and transit time. As can be seen, the obtained results are pretty similar to those presented previously in Figure 5.11. However, with feature reduction, the model complexity was reduced.



(a) Packet delivery ratio.

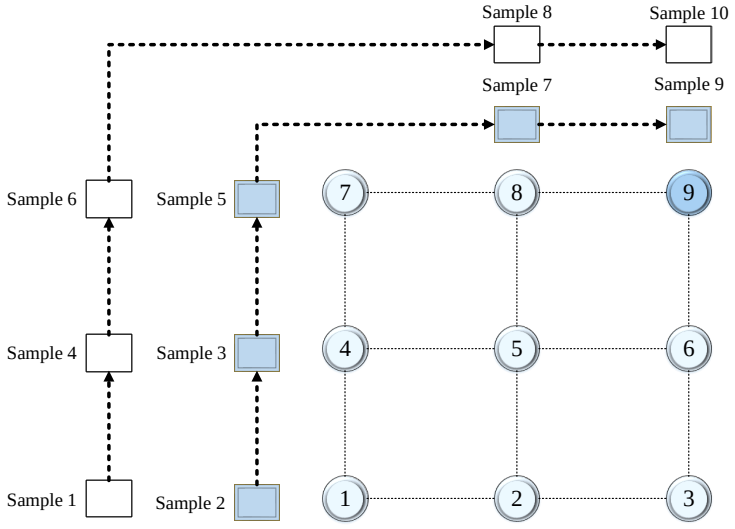
(b) Network attained throughput

(c) Network transit time.

FIGURE 5.15: Packet delivery ratio, network throughput and transit time (selected features are computed for a target accuracy of 85 %).

5.3 Machine learning based on traffic priorities

In the previous section, a machine learning congestion control mechanism was presented. However, the proposed classifier does not take into account the priority of the traffic class. Similar to the solutions presented in previous chapters, each traffic type has its own QoS requirements where the traffic class 1 has the highest priority and traffic 4 has the lowest priority. Therefore, a new dataset that includes different types of traffic has to be built. For this purpose, the data collection and data processing steps must be modified. In order to understand the methodology used to create a dataset with different traffic types, a simple example will be used. This approximation is shown in Figure 5.16 for two packets that belong to two different traffic classes.



Repository

sample	class	Packet ID	ρ	q	time
1	1	1	ρ_1	q_1	t_1
2	2	2	ρ_1	q_1	t_2
3	2	2	ρ_4	q_4	t_3
4	1	1	ρ_4	q_4	t_4
5	2	2	ρ_7	q_7	t_5
6	1	1	ρ_7	q_7	t_6
7	2	2	ρ_8	q_8	t_7
8	2	2	ρ_8	q_8	t_8
9	2	2	ρ_9	0	t_9
10	1	1	ρ_9	0	t_{10}

FIGURE 5.16: Passive monitoring technique used to build a training dataset considering different traffic types.

For the new data collection process, the traffic flows were implemented as follows. The two packets are transmitted at the same time, and they have a *commonID* identifier included in their packet header that links them hop by hop. Basically, the objective of having a common identifier *commonID*, is to analyze how the transmission of data belonging to traffic type 2 penalize the traffic type 1 in terms of successful packet reception and network packet transit time. Besides, in this example the two traffic type are transmitted at the same rate. However, in the real data collection used later, different combinations of traffic patterns were implemented. For instance, one combination is maintaining the rate of traffic type 1, and increasing the rate of traffic type 2, and observing how the traffic type 1 reduces its performance.

In order to build the training dataset, the network transit time requirements for each class are verified according to the example values presented in Table 5.5. First, if the network packet transit time requirement for traffic type 1 is not met, the packet is considered as "not received". Thus, the traffic types with lower priorities are set up to not received. Remember that this is an approximation, and maybe the transmission of the packet traffic type 2 was not the only cause for the traffic type 1 packet loss. Second, the same approach will be applied when transmitting more than two traffics. Remember that all the data processing is done thanks to the *commonId* identifier which links the different traffic types.

TABLE 5.5: Network transit time values used in the data processing.

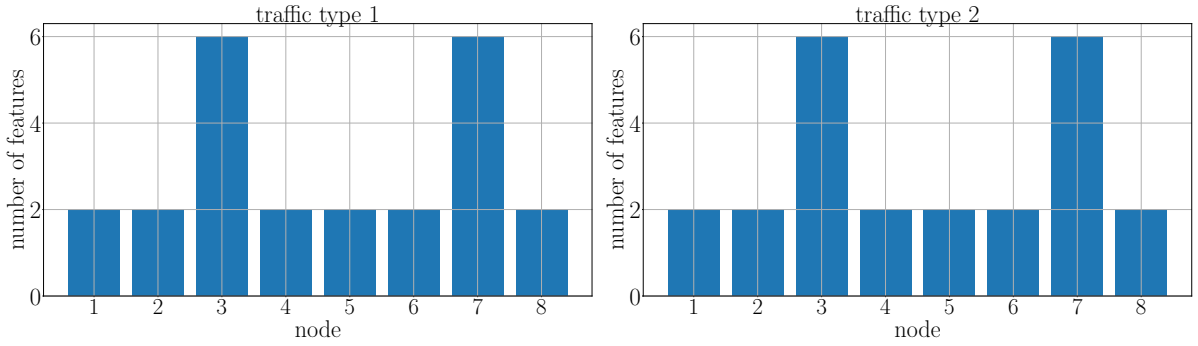
Class	Network transit time
1	50 ms
2	100 ms

In the next subsection we will evaluate a Smart Grid data scenario where two types of traffic are transmitted. On the one hand, the system will provide congestion control in situations of high network load. On the other hand, the solution will give higher transmission priority to traffic type 1.

5.3.1 Simulation results

For the evaluation of the mechanism, the same previous NAN scenario has been considered. Besides, there are two traffic types transmitted upstream from the all smart meters towards the data concentrator. The packet length and the packet generation rate distributions have been selected as exponential for both data flows, and their values are 200 Bytes and 100 pkt/s respectively.

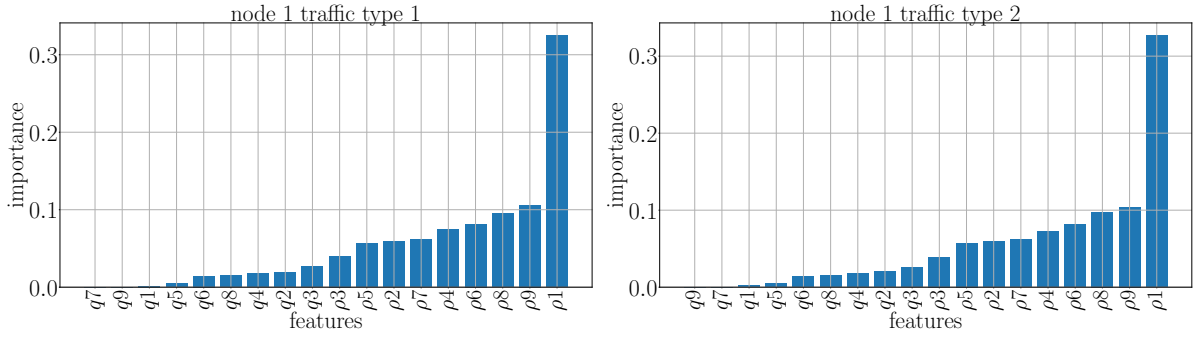
In order to build the training dataset, feature selection was done for a target accuracy of 85%. As previously said, with this the complexity (number of decision rules) of the machine learning model is reduced. Figure 5.17 shows the number of features needed to build the model for node and for traffic type. From the plots, it can be seen how the number of features has been reduced from a maximum value of 18 features to a value of 2 for most of the nodes.



(a) traffic type 1. (b) traffic type 2.

FIGURE 5.17: Number of features needed to target an accuracy score 85% per node and per traffic. Network size 9 nodes

Keep in mind that the dataset was split per node and per traffic, and so the samples that were not generated for the node itself are discriminated in the training. Therefore, each node will have its own decision rules based on the traffic type to be transmitted. Figure 5.18 shows the most relevant features for node 1 and for the two traffics types. As can be seen, the most important features are the networks parameters which belong to each node itself. Furthermore, the most relevant feature is too high compared with the less important features. Therefore, it is not needed to use all the features for training the model as it will be shown in the results.

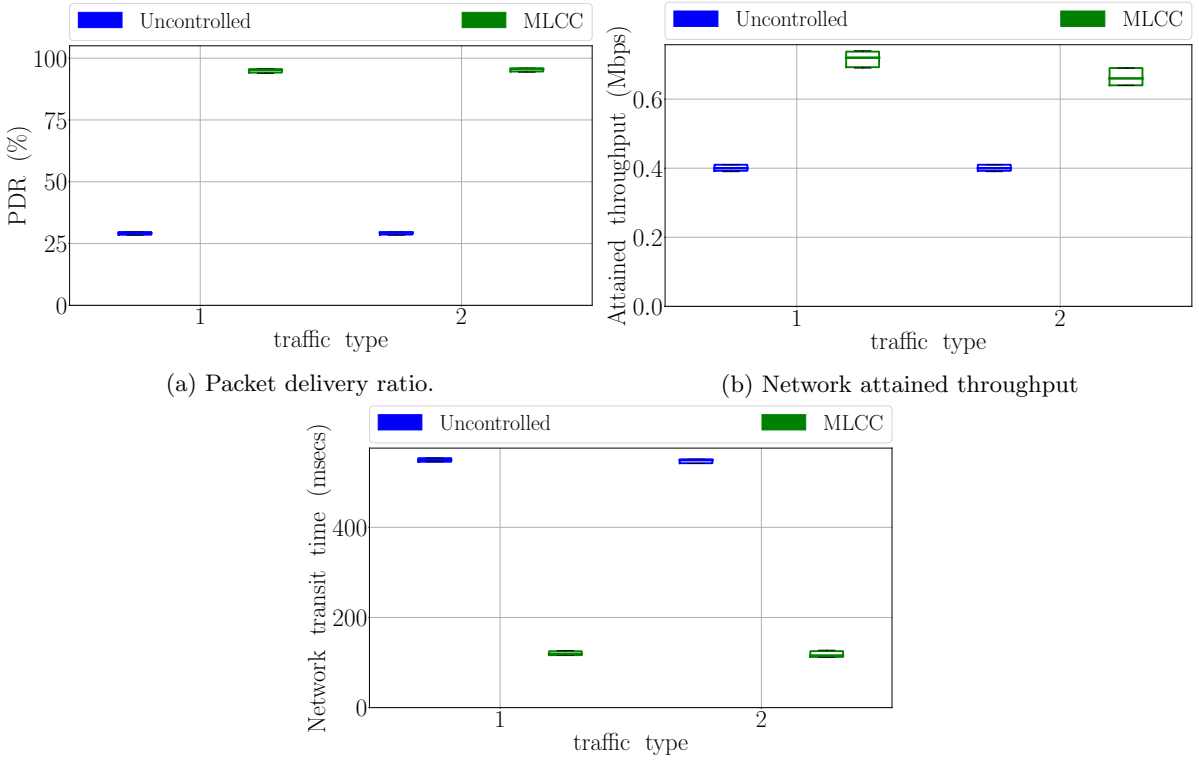


(a) traffic type 1. (b) traffic type 2.

FIGURE 5.18: Feature importance for node 1. Network size 9 nodes

When training the model, the decision tree classifier provided a higher prediction power when the selected tree depth is 6, and the chosen criterion is entropy. As previously said, the decision tree parameters were chosen after doing performing several evaluations with different combinations of parameters.

Figure 5.19 shows the obtained results in terms of packet delivery ratio, network throughput and transit time. Figure 5.19a shows that in the absence of a congestion control solution a high percentage of packets will be lost on their way towards the data concentrator. However, when the ML is applied a 92% of the packers are correctly transmitted to the sink node. Regarding the attained throughput, as can be seen in Figure 5.19b the value or traffic type 1 is higher than traffic class 2, and thus, different priorities are correctly given to each traffic. Besides, the attained throughput is higher compared to the uncontrolled case for both traffics when the proposed solution is applied. Finally, Figure 5.19c shows the significant improvements obtained in terms of network transit time with our solution.



(a) Packet delivery ratio.

(b) Network attained throughput

(c) Network transit time.

FIGURE 5.19: Packet delivery ratio, network throughput and transit time (selected features are computed for a target accuracy of 85 %).

Overall, the previous results indicate the benefits obtained from our congestion control mechanism when two traffics are transmitted on the network. Next, a last set of experiments will be carried out to verify if the approximation used to differentiate traffic types works for four traffics types and a larger scenario. In a similar way, the traffic type 1 has the highest QoS needs, while the traffic type 4 has the lowest QoS requirements. For this purpose, a new data collection and data processing have to be done as follows. First, if the QoS requirements of traffic type 1 are not met (packet is not received correctly or out of time ranges), the lower priority traffics are not considered for transmission. The same will happen if the QoS needs of traffic 2 are not met. That is traffic type 3 and 4 will not be considered for transmission. As before, for the

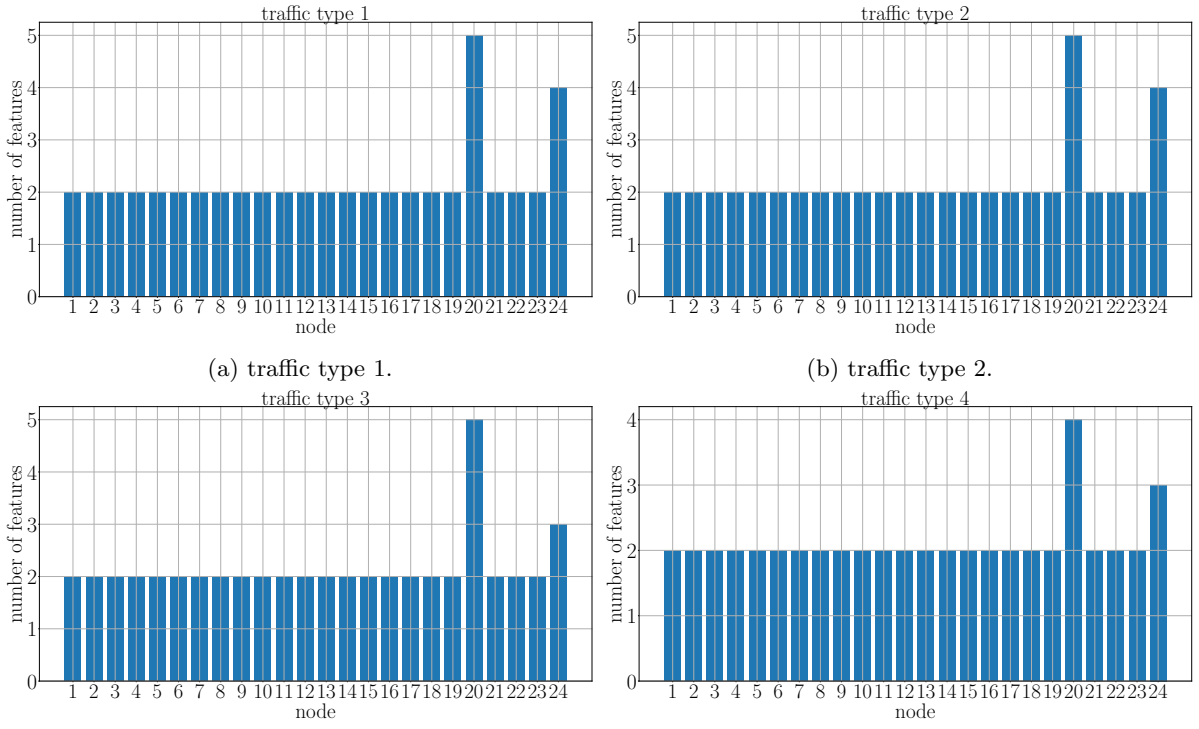
data collection a *commonID* identifier is included in the packet header in order to link the four traffic types. The network transit time requirements that will be used for data processing are presented in Table 5.6. Besides, different combinations of traffic patters were done in the data collection process.

TABLE 5.6: Network transit time values used in the data processing.

Class	Network transit time
1	50 ms
2	100 ms
3	1000 ms
4	2000 ms

For the evaluation, a network size of 25 nodes has been considered where there are 24 smart meters transmitting data towards the data concentrator. The packet length and the packet generation rate distributions have been selected as exponential for the four data flows, and their values are 200 Bytes and 20 pkt/s respectively.

Given that the current evaluation scenario consists of 25 nodes, the training dataset will contains 50 features. In order to reduce the dataset dimension as well as the model complexity, a feature selection is performed again for a target accuracy of 85%. Figure 5.20 shows how the number of features needed to generate the learning model is reduced from 50 to almost 2 features for almost all the nodes. In this case, feature selection greatly reduced the model complexity. With this strategy, each node will train its model based on its most relevant features as it is shown in Figure 5.21. For this case, the classifier predictive power performed better when the tree depth is configured to 4.



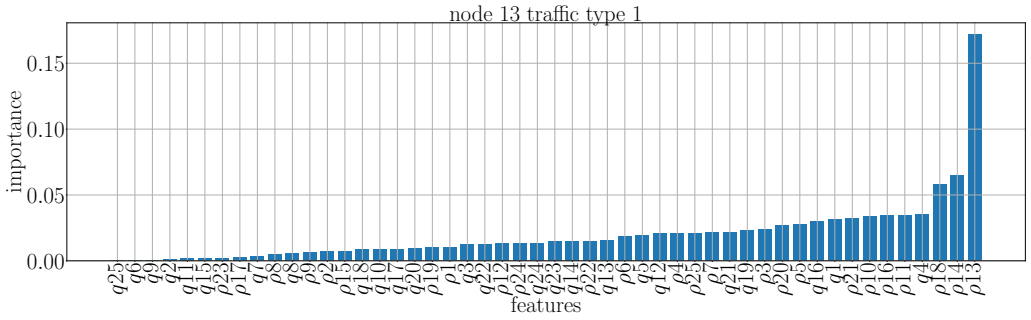
(a) traffic type 1.

(b) traffic type 2.

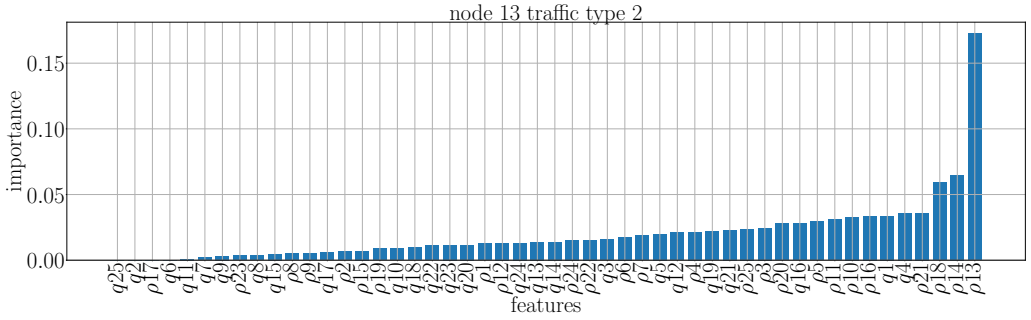
(c) traffic type 3.

(d) traffic type 4.

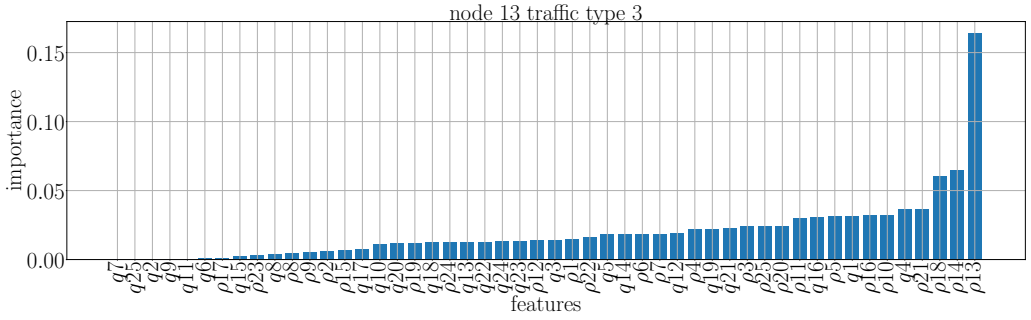
FIGURE 5.20: Number of features needed to reach the target accuracy score of 85%. Network size 25 nodes).



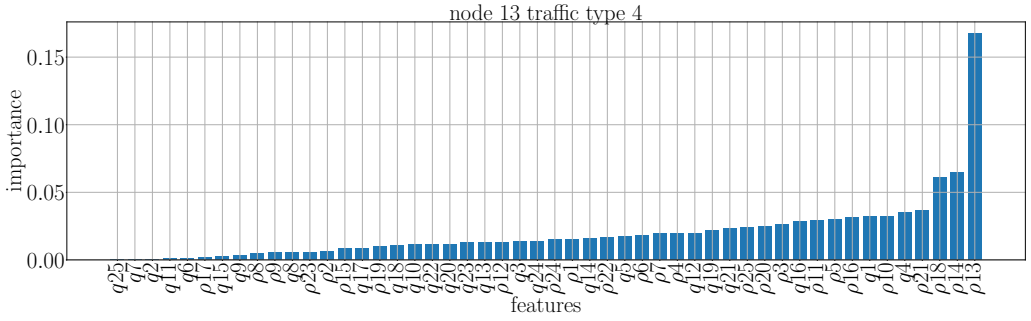
(a) traffic type 1.



(b) traffic type 2.



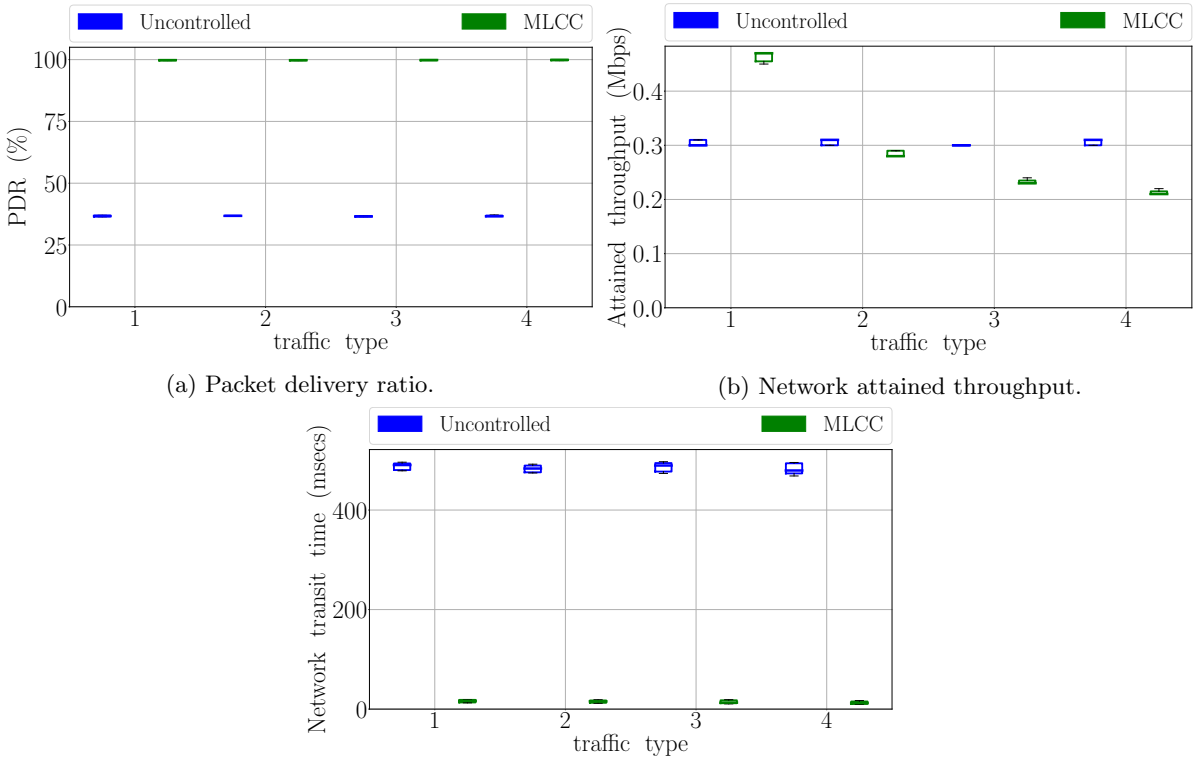
(c) traffic type 3.



(d) traffic type 4.

FIGURE 5.21: Feature importance for node 13. Network size 25 nodes.

The obtained results are shown in Figure 5.22. Figure 5.22a shows that in the absence of a congestion control solution a 75% of packets will be lost on their way towards the data concentrator. However, with the proposed solution, the PDR is 100 % for all traffic types. Besides, Figure 5.22b shows how in network congestions situations the traffic type 1 is strongly favoured in terms of throughput, and different priorities are given for each traffic type. Finally, Figure 5.22c shows the significant improvements obtained in term of network transit time with our solution.



(a) Packet delivery ratio.

(b) Network attained throughput.

(c) Network transit time.

FIGURE 5.22: Packet delivery ratio, network throughput and transit time (selected features are computed for a target accuracy of 85 %).

5.4 Conclusions and future work

In this chapter, a congestion control mechanism based on machine learning has been presented, implemented and evaluated from scratch.

First, a data collection process that involves different network traces have been implemented. These network allow to capture valuable network parameters at network, MAC and physical layers. In other words, to capture all the traffic going through the network. There are different NAN applications data packets transmitted over the network, and taking advantage of this, a learning model will be generated to make knowledge from historical data. With this strategy, the proposed mechanism will predict whether or not to transmit the NAN applications based on the current network load. The forwarding decision rules depend on the traffic type and its QoS needs.

As mentioned throughout the chapter, the predictive power of the model is based on how meaningful is the training dataset. Thus, in the first part of the chapter how to correctly generate the training dataset was explained. It was concluded that disaggregating the training dataset per node and then per traffic increases the prediction power of the self-learning model. Besides, feature reduction was implemented to reduce the complexity of the model. For this purpose, L1 regularization was computed to select the most meaningful features.

The decision tree classifier was used to generate the prediction model. DTs are easy to implement, and for our case, gives a good performance in terms of accuracy. However, other learning models could have been implemented as well. The classifier performance lies in the dataset itself

and model parameters used for training. In order to evaluate the classifier, the ROC curve was used to assess the prediction power to unseen data.

In addition, the models was also evaluated again in a Smart Grid NAN scenario. The obtained results show how the proposed solution improves the network performance in terms of PDR, network throughput and transit time.

It should be noted that this machine learning based congestion control mechanism is based on a programmed framework. This framework is responsible for generating the network traces, processing the captured traffic, training the models, and finally exporting them in the simulator.

As future lines of work, machine learning signaling will be proposed and implemented in order to propagate the network parameters to the whole networks. In this work, it has been assumed that all the nodes knows the network state at anytime. However, it is an ideal case and a robust control signaling has to be added. Besides, other features such as emergency awareness and fairness in the distribution of network resources among all the network nodes, will also be considered.

Chapter 6

Conclusions and Future Work

Wireless mesh networks (WMN) have attracted the attention of many researchers for a long time. They have been evaluated in many research areas and one of them is the Smart Grid networks. Smart grids are an evolution for the traditional electricity networks, where a data communication network is available to complement the traditional electrical infrastructure. On the one hand, the electricity network is responsible for generating, transporting and distributing this valuable resource to subscribers. On the other hand, the data communication network includes a robust architecture that provides a large number of services. The Smart Grid data communication network comprises three subnetworks: the Home Area Network (HAN), the Neighborhood Area Network (NAN) and the Wide Area Network (WAN). Each part of the Smart Grid data communication network fulfills different purposes.

In this thesis, the proposed solutions are focused on improving the network performance for the Neighborhood Area Networks. These networks interconnect the different smart meters and other home appliances with a data concentrator. Several NAN applications are transmitted bidirectionally. For instance, meter reading, power quality or electric vehicle traffic charge information are transmitted from the smart meters towards the data concentrator. Meanwhile, the control center transmits for instance demand response information in order to take corrective solutions in anomalous network situations.

Among the different network technologies available nowadays, Wireless mesh networks have emerged as a very promising candidate to improve the utility, performance and usability of the NANs. Thus, the main contributions of this thesis are devoted to improve the data routing, provide fair distribution of the network resources, and finally to present different solutions to deal with network congestion. The following subsections summarize the most important characteristics of each of the proposals, as well as some future lines of work.

6.1 Multi-path and multi-channel routing for HWMP

The first contribution of this thesis is the proposal of a new routing mechanism for the Smart Grid NANs when the selected network technology is the IEEE 802.11s WMN, the Multi-Path Multi-Channel Hybrid Wireless Mesh Protocol (MPC-HWMP), which is a modification of the default HWMP. MPC-HWMP computes all the possible paths between the smart meters and the data concentrator. With this strategy, the most priority traffics are transmitted through the best paths, while the traffic types with lower QoS needs are transmitted for the worst paths.

Keep in mind that the paths are ranked according to the airtime link metric. That is, the best paths are the ones with the lowest metric. In order to improve even more the benefits obtained, a multichannel allocation mechanism was also proposed. The multichannel allocation scheme separates the control and the data traffic in different channels. The dedicated control channel is assigned for the path discovery mechanism. With the multipath proposal, there are more control frames transmitted in order to maintain the multiple paths, and thus, the protocol overhead is increased. In order to mitigate this problem, the data channels are used exclusively to transmit the NAN applications. Besides, the most priority traffic types are transmitted for the less congested channels. Basically, the proposed solution allows a better distribution of network resources, since all smart meters transmit a portion of the network traffic.

In order to evaluate this proposal, an Smart Grid NAN scenario with different network sizes and bidirectional data traffic was considered. Besides, the default HWMP and MPC-HWMP proposal were evaluated for two different load conditions: a high network load and an extreme congestion situation. The numerical results showed how as the network size is increased, the network became more congested. Therefore, PDR and network throughput values are decreased for all traffic types. However, the attained throughput and PDR are higher when the MPC-HWMP was used, and the benefits obtained are even greater in the extreme congestion situation. Besides, each traffic type is served based on its priority (QoS). When evaluating the network transit time, the best performance is again obtained with the MPC-HWMP proposal.

When using the MPC-HWMP proposal, the routing table size is increased in order to store the multiple paths to a given destination. Therefore, the cost in terms of memory needed (MBytes) to store the multiple entries was also evaluated. It was concluded that the increased memory is not a problem in a NAN environment, where the number of nodes is fixed and they could not grow indefinitely. Furthermore, these additional memory requirements are not a problem for the current devices. On the other hand, the number of PREQ and PREP messages transmitted over the NAN are also increased with MPC-HWMP. Thus, the protocol overhead was also evaluated. For this purpose, the control channel utilization factor was also analyzed. It was observed that the control channel utilization factor never reaches high values.

Overall, the results obtained showed the improvements in terms of PDR, network throughput and transit time. As future lines of works, nodes reputation will be added to the multipath routing mechanism. For this purpose, graph theory and conflicts graphs will be used.

6.2 Emergency and fairness aware mechanisms for congestion control

As previously stated, there are several applications transmitted over the NAN, and some of them can be affected for network congestion and emergencies situations. In this context, traffic types with higher QoS needs have to be prioritized because most of them are critically important for the correct NAN operation. In this chapter, the improvements obtained by means of the application of two new congestion control proposals are summarized.

The fist mechanism is the Emergency Aware Congestion Control - HWMP (EA-HWMP) which deals with network congestion and emergency situations. The selected technology is again the wireless mesh networks made up of smart meters and one data concentrator. The proposed mechanism is based on the definition of some congestion control functions, which provide different transmission probabilities taking into account the current network load, the emergency

state and the traffic type. To this end, each node periodically measures the current channel utilization factor and the emergency state and, based on these measurements, different forwarding decision are taken. When the network is congested, the mechanism increases the transmission probability for priority traffics by discarding less important traffics. Besides, the transmission probability is even higher when the emergency state is increased. Emergency states are intrinsic or extrinsic to the NAN, and they can be triggered manually or automatically. To notify the emergency states, special messages are broadcast to the whole network by using a dedicated control channel.

Furthermore, a multichannel allocation scheme was coupled to the proposed solution. In the same way, the multichannel scheme involved two tasks: split data and control traffic, and channel switching. In a not congested scenario, NAN applications are transmitted only for one data channel. The ideal is to transmit all data by using one channel in order to avoid the problems related to channel switching. However, if the current channel exceeds a certain congestion threshold, more channels are used for data transmission to alleviate the network.

The EA-HWMP was evaluated again in a Smart Grid NAN environment where different network sizes were considered. This proposal was compared with the basic HWMP in two scenarios. The first scenario tested how the congestion control mechanism reacts when the most priority traffic type begins to saturate the network. In this case, the congestion control functions based on the local measurements prioritize this traffic class by discarding less priority traffic types. The second scenario contemplated three different emergency states (normal, medium and high) with high network load. From the obtained results, it can be seen how the system adjusted the congestion control functions in order to prioritize again the most priority traffic types, specially in high emergency situations.

Similar to the previous proposal, EA-HWMP and HWMP were evaluated in terms of PDR, delivered throughput, transit time and channel utilization factor. The result showed the benefits obtained when the EA-HWMP is used, specially in high network loads and in high emergency states.

The second proposed congestion control mechanism has the main objective of obtaining a fair distribution of network resources to all the network nodes. Note that smart meters are static nodes and therefore, some of them may be more favored depending on their geographical location. That is, nearer nodes to the data concentrator could be able to deliver higher throughput to the data concentrator. Besides, some nodes can monopolize the use of network resources if their packet generation rates are not regulated. Therefore, a feedback has to be given to the sources in order to not overload the network. In this context, the proposed Fair and Distributed Congestion Control (FDCC) consists of a set of algorithms to deal with network congestion and fairness in the QoS provision. With FDCC, all the nodes have the same opportunities to correctly deliver their NAN applications to the data concentrator regardless of their geographical location and transmission rate. Furthermore, the distributed solution is applied in the source nodes and not in the relay nodes as was presented in the previous proposals. The objective is mainly not to transmit packets that will be dropped on their way to the data concentrator.

FDCC was evaluated in terms of PDR, network throughput, network transit time and fairness metrics. Besides, FDCC was compared with the EDCA mechanism, and the benefits are remarked with the proposed mechanism. Basically, FDCC and EDCA mechanisms provides traffic differentiation in different ways. When EDCA is used, a full priority of transmission is given to a single class and it is applied locally. That is, EDCA processes frames according only to their traffic types and does not differentiate the flows from the different source nodes. Therefore, EDCA mechanism is independent of the global state of the network. However, the FDCC

solution is applied globally and all source nodes and their data flows are fair served according to its priority.

The obtained results showed how the network congestion is avoided with FDCC (PDR of almost 100%), and all the network resources are fair distributed to all network nodes (same attained throughput). Besides, all traffic types met their minimum QoS requirements.

It should be emphasized that the algorithm is agnostic to the routing protocol, network and MAC layers. In this case, and to carry out and provide studies on a greater diversity of scenarios, a classical wireless ad hoc network was the selected technology to deploy the NAN. The AODV routing protocol and the latest IEEE 802.11ac were used in the network and physical layer respectively.

As future work, the congestion control mechanism will be applied in different environments, specially in crowded scenarios with mobile nodes. Besides, an emergency system will be coupled to the FDCC proposal.

6.3 Machine learning congestion control mechanism

Today there is a large amount of data available, and many researchers have focused their efforts on generating knowledge from this data. With this in mind, in this thesis a congestion control mechanism based on machine learning techniques has also been proposed. These decisions are based also on the traffic type. That is, priority traffic types will be more likely to be transmitted.

Data collection is the most important process to generate the model. For this reason, in this thesis a framework was programmed to create the model from scratch. That is, the framework generates different network traces to capture all the traffic transmitted on the network. These traces represents different network parameters such as channel utilization factor or buffer occupancy. Furthermore, the methodology used to create a dataset that contemplates traffic differentiation is deeply covered. Besides, how to process this data and generate a good and meaningful data set was also proposed. To this end, techniques such as feature reduction and granularity of the dataset were employed. These techniques allow to reduce the model complexity (number of decision rules), and also to increase the prediction power to unseen data. Therefore, the most important features are used for training as well as dividing the dataset according to the belonging characteristics of the node. That is, samples that were not generated by the node are discriminated in the learning process.

The learning model selected is the decision trees, and their best tuning parameters were chosen though an extensive number of simulations. This classifier was evaluated with the well-known performance metric ROC curve. The validation exhibited the high predictive power of the model generated. The prediction power was also evaluated in a more realistic environment through network simulations. For this purpose, a NAN scenario was evaluated for different cases. Basically, in each case, the number of traffics types were increased. In the first case, just one traffic type was considered, and it was used to evaluate the congestion control mechanism. The other two cases considered a larger network size where more traffic classes are transmitted. Here, the congestion control mechanism and traffic differentiation were evaluated together. The obtained results showed how the network performance (PDR, network throughput and transit time) is increased when the machine learning model is used. Besides, the trained model is able to provide traffic differentiation based on the current network load. Finally, the methodology

carried out for data collection with different traffic priorities performed well. However, as future work other methodologies will be proposed.

As a final conclusion it can be emphasized that all the thesis objectives were addressed, and the most important characteristics of each proposal are summarized in Table 6.1:

TABLE 6.1: Comparison among the different proposal of this dissertation.

Solution	Features	Benefits obtained	Cost
MPC-HWMP	- Multi-path routing. - Multi-channel allocation. - Paths are ranked based on the airtime link metric and the best data channels are ranked based on the channel utilization factor.	- No need for metric modification. - Better distribution of network resources, priority traffics are transmitted over the best paths and less congested channels. - Control and data traffic are split in different channels. The control channel is used for the path discovery mechanism, while the data channels are used to transmit the data traffic. - Forwarding decision rules are made hop by hop, so, routing loops can appear. Routing loops are solved with a hop count strategy.	- Routing table size is increased. - Additional control traffic to maintain the multiple paths. - More energy used for the multichannel switching. - Performance improvements are strongly related to the number of available paths, and even more to the number of available data channels.
EA-HWMP	- Congestion control mechanism is based on congestion control functions. - Possible emergency states are considered. - Emergency system is based on network symptoms. - Multi-channel allocation.	- Transmission probability varies depends on the current network load. - The packet (re)transmission is made by using the minimum possible number of data channels. - Different congestions control functions are defined for each traffic type and emergency situation. - Control and data traffic are split in different channels. The control channel is used for the path discovery mechanism and the emergency signalling, while the data channels are used to transmit the data traffic.	- Congestion control functions are fixed. - Additional control traffic to maintain emergency system module. - More energy used for the multichannel switching. - Forwarding decision rules are made hop by hop. Packets are likely to be discarded in the relay nodes. - Performance improvements are strongly related to the number of available data channels.
FDCC	- Fairness of the distribution of the network resources. - Congestion control mechanism based on set of algorithms the runs individually at each node. - Traffic differentiation based on the channel utilization factor.	- All nodes have the same possibility to transmit their data to data concentrator regardless their geographical position or their packet generation rates (unbalanced source rates). - Forwarding decision rules are made in the source nodes. Packets likely to be discarded in the relay nodes are not transmitted. - Solution is agnostic the network and mac layers.	- Additional signaling in order to adjust the packet generation rates. - Tuning algorithms parameters is mandatory, and bad tuning can lead to a poor performance.
MLCC	- Congestion control mechanism based on learning rules. - Traffic differentiation.	- Prediction of the packet network transit time. - Forwarding decision rules are made in the source nodes. Packets likely to be discarded in the relay nodes are not transmitted. - Others features can be easily included (emergency state awareness and fairness)	- Data collection is mandatory, and difficult to perform in some cases. - Data processing (feature extraction). - Feature selection can lead to complex signaling.

Bibliography

[1] Raschka Sebastian and Mirjalili Vahid. *Python Machine Learning Machine Learning and Deep Learning with Python scikit-learn and TensorFlow 2*. Packt Publishing Ltd, 2019.

[2] NS-3 Consortium. Ns-3 is a discrete-event network simulator. "https://www.nsnam.org/".

[3] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.

[4] Yu Liu, Kin-Fai Tong, Xiangdong Qiu, Ying Liu, and Xuyang Ding. Wireless mesh networks in iot networks. In *2017 International Workshop on Electromagnetics: Applications and Student Innovation Competition*, pages 183–185. IEEE, 2017.

[5] Agrawal; D.; & Zeng; Q. A. *Introduction to wireless and mobile systems*. Cengage learning, 2015.

[6] Achieving Load Balancing in Wireless Mesh Networks Through Multiple Gateways. In *2006 IEEE International Conference on Mobile Ad Hoc and Sensor Sysetems*, pages 807–812. IEEE, IEEE, oct 2006. ISBN 1-4244-0506-8. doi: 10.1109/MOBHOC.2006.278655. URL http://ieeexplore.ieee.org/document/4054001/.

[7] Szymon Szott. Selfish insider attacks in IEEE 802.11s wireless mesh networks. *IEEE Communications Magazine*, 52(6):227–233, jun 2014. ISSN 0163-6804. doi: 10.1109/MCOM.2014.6829968. URL http://ieeexplore.ieee.org/document/6829968/.

[8] Xiaoheng Deng, Lifang He, Qiang Liu, Xu Li, Lin Cai, and Zhigang Chen. EPTR: expected path throughput based routing protocol for wireless mesh network. *Wireless Networks*, 22(3):839–854, apr 2016. ISSN 1022-0038. doi: 10.1007/s11276-015-1003-3. URL http://link.springer.com/10.1007/s11276-015-1003-3.

[9] Md Asri Bin Ngadi, Saqib Ali, Abdul Hanan Abdullah, and Rashid Hafeez Khokhar. A taxonomy of cross layer routing metrics for wireless mesh networks. *EURASIP Journal on Wireless Communications and Networking*, 2012(1):177, 2012. ISSN 1687-1499. doi: 10.1186/1687-1499-2012-177. URL http://jwcn.eurasipjournals.springeropen.com/articles/10.1186/1687-1499-2012-177.

[10] Jing Chen, Kun He, Ruiying Du, Minghui Zheng, Yang Xiang, and Quan Yuan. Dominating Set and Network Coding-Based Routing in Wireless Mesh Networks. *IEEE Transactions on Parallel and Distributed Systems*, 26(2):423–433, feb 2015. ISSN 1045-9219. doi: 10.1109/TPDS.2013.303. URL http://ieeexplore.ieee.org/document/6678512/.

[11] Wenxiao Shi, Shuo Shang, Yu Zheng, and Yinlong Xu. Routing Metric of Expected Delay in Multi-Radio Multi-Channel Wireless Mesh Networks. *Journal of Communications*, 9 (11):851–858, 2014. ISSN 23744367. doi: 10.12720/jcm.9.11.851-858.

[12] Anfu Zhou, Min Liu, Zhongcheng Li, and Eryk Dutkiewicz. Cross-layer design with optimal dynamic gateway selection for wireless mesh networks. *Computer Communications*, 55:69–79, 2015. ISSN 01403664. doi: 10.1016/j.comcom.2014.08.011. URL http://dx.doi.org/10.1016/j.comcom.2014.08.011.

[13] Jatinder Singh Saini and Balwinder Singh Sohi. A Survey on Channel Assignment Techniques of Multi-Radio Multi-channel Wireless Mesh Network. *Indian Journal of Science and Technology*, 9(42), nov 2016. ISSN 0974-5645. doi: 10.17485/ijst/2016/v9i42/92192. URL http://www.indjst.org/index.php/indjst/article/view/92192.

[14] Somayeh Kafaie, Yuanzhu Chen, Octavia A Dobre, and Mohamed Hossam Ahmed. Joint inter-flow network coding and opportunistic routing in multi-hop wireless mesh networks: A comprehensive survey. *IEEE Communications Surveys & Tutorials*, 20(2):1014–1035, 2018.

[15] Amal A. Eltahir, Rashid A. Saeed, Amitava Mukherjee, and Mohammad Kamrul Hasan. Evaluation and analysis of an enhanced hybrid wireless mesh protocol for vehicular ad hoc network. *EURASIP Journal on Wireless Communications and Networking*, 2016 (1):169, dec 2016. ISSN 1687-1499. doi: 10.1186/s13638-016-0666-5. URL https://jwcn-eurasipjournals.springeropen.com/articles/10.1186/s13638-016-0666-5.

[16] Rasmus Liborius Bruun and Troels Bundgaard Sørensen. Analysis of link loss in wireless mesh networks based on field trials in disaster communication scenarios. In *2019 IEEE 30th Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, pages 1–6. IEEE, 2019.

[17] Mehmet Ali Ertürk, Muhammed Ali Aydin, Luca Vollero, and Roberto Setola. Ieee 802.11 s mesh network analysis for post disaster communication. In *International Telecommunications Conference*, pages 53–59. Springer, 2019.

[18] Gudi Siva Leela Krishna Chand, Manhee Lee, Soo Young Shin, et al. Drone based wireless mesh network for disaster/military environment. *Journal of Computer and Communications*, 6(04):44, 2018.

[19] Mariusz Wzorek, Cyrille Berger, and Patrick Doherty. Router node placement in wireless mesh networks for emergency rescue scenarios. In *Pacific Rim International Conference on Artificial Intelligence*, pages 496–509. Springer, 2019.

[20] Suresh Sankaranarayanan. Hierarchical intelligent agent based wireless body sensor mesh networks. In *2009 Sixth International Conference on Information Technology: New Generations*, pages 1602–1603. IEEE, 2009. URL https://ieeexplore.ieee.org/abstract/document/5070867.

[21] Kashif Saleem, Khan Zeb, Abdelouhid Derhab, Haider Abbas, Jalal Al-Muhtadi, Mehmet A Orgun, and Amjad Gawanmeh. Survey on cybersecurity issues in wireless mesh networks based ehealthcare. In *2016 IEEE 18th International Conference on e-Health Networking, Applications and Services (Healthcom)*, pages 1–7. IEEE, 2016.

[22] Wanqing Tu, Cormac J Sreenan, Sanjay Jha, and Qian Zhang. Multi-source video multicast in internet-connected wireless mesh networks. *IEEE Transactions on Mobile*

Computing, 16(12):3431–3444, 2017. URL https://ieeexplore.ieee.org/abstract/document/7893751.

[23] Muni Lavanya, C Shoba Bindu, and G Vijay Kumar. Internet traffic based channel selection in multi-radio multi-channel wireless mesh networks. *International Journal of Communication Networks and Information Security*, 11(2):262–269, 2019. URL https://ieeexplore.ieee.org/abstract/document/7868436.

[24] Andres Marcelo Vazquez Rodas. *Contribution to the Improvement of the Performance of Wireless Mesh Networks Providing Real Time Services*. PhD thesis, 2015.

[25] Silvio Sampaio, Pedro Souto, and Francisco Vasques. A review of scalability and topological stability issues in ieee 802.11 s wireless mesh networks deployments. *International Journal of Communication Systems*, 29(4):671–693, 2016.

[26] Bart Braem, Chris Blondia, Christoph Barz, Henning Rogge, Felix Freitag, Leandro Navarro, Joseph Bonicioli, Stavros Papathanasiou, Pau Escrich, Roger Baig Viñas, and Others. A case for research with and on community networks. *ACM SIGCOMM Computer Communication Review*, 43(3):68–73, jul 2013. ISSN 01464833. doi: 10.1145/2500098.2500108. URL http://dl.acm.org/citation.cfm?doid=2500098.2500108.

[27] José Maria Saldaña Medina, Andres Arcia-Moret, Ermanno Pietrosemoli, Marco Zennaro, Bart Braem, and Arjuna Sathiaseelan. Alternative Network Deployments: Taxonomy, Characterization, Technologies, and Architectures. Technical report, 2016.

[28] freinfunk.net. Freifunk is a non-commercial initiative for free wireless networks. "https://freifunk.net/en/", .

[29] freinfunk.net. Athens wireless metropolotan network. "http://www.awmn.net/", .

[30] FunkFeuer. FunkFeuer Free net. "https://www.funkfeuer.at/".

[31] guifi.net. guifi.net. "https://guifi.net/".

[32] Ninux. Ninux. "http://ninux.org/".

[33] W. Meng, R. Ma, and H. Chen. Smart grid neighborhood area networks: a survey. *IEEE Network*, 28(1):24–32, January 2014. ISSN 0890-8044. doi: 10.1109/MNET.2014.6724103.

[34] Shengjie Xu, Yi Qian, and Rose Qingyang Hu. On Reliability of Smart Grid Neighborhood Area Networks. *IEEE Access*, 3:2352–2365, 2015. ISSN 2169-3536. doi: 10.1109/ACCESS.2015.2502250. URL http://ieeexplore.ieee.org/document/7332240/.

[35] Murat Kuzlu, Manisa Pipattanasomporn, and Saifur Rahman. Communication network requirements for major smart grid applications in han, nan and wan. *Computer Networks*, 67:74–88, 2014.

[36] Yasin Kabalci. A survey on smart metering and smart grid communication. *Renewable and Sustainable Energy Reviews*, 57:302–318, 2016.

[37] Manali Gupta, Shirshu Varma, et al. Configuration of aerial mesh networks with internet of things. In *2018 International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET)*, pages 1–3. IEEE, 2018.

[38] Jonghun Kim and Jaiyong Lee. Cluster-based mobility supporting wmn for iot networks. In *2012 IEEE International Conference on Green Computing and Communications*, pages 700–703. IEEE, 2012. URL https://ieeexplore.ieee.org/abstract/document/6468394.

[39] Xin Jiang, Mingzhe Liu, Chen Yang, Yanhua Liu, and Ruili Wang. A blockchain-based authentication protocol for wlan mesh security access. *CMC-Comput Mater Continua*, 58 (1):45–59, 2019.

[40] Alberto Attilio Brincat, Alfio Lombardo, Giacomo Morabito, and Salvatore Quattropani. On the use of blockchain technologies in wifi networks. *Computer Networks*, 162: 106855, 2019. doi: https://doi.org/10.1016/j.comnet.2019.07.011. URL http://www.sciencedirect.com/science/article/pii/S1389128619306073.

[41] Chenlei Pan, Bo Liu, Haibo Zhou, and Lin Gui. Multi-path routing for video streaming in multi-radio multi-channel wireless mesh networks. In *2016 IEEE International Conference on Communications (ICC)*, pages 1–6. IEEE, 2016. URL https://ieeexplore.ieee.org/abstract/document/7511222.

[42] Jerrin Sebastian, Deepu Job, and Smitha Jacob. Qos based congestion control algorithm for video traffic in wireless mesh network. *International Research Journal of Engineering and Technology (IRJET)*, 3(08):363–368, 2016.

[43] Adriana Hava, Yacine Ghamri-Doudane, John Murphy, and Gabriel-Miro Muntean. A load balancing solution for improving video quality in loaded wireless network conditions. *IEEE Transactions on Broadcasting*, 65(4):742–754, 2019.

[44] Mohammad Tariq Meeran, Paul Annus, Muhammad Mahtab Alam, and Yannick Le Moullec. Evaluation of voip qos performance in wireless mesh networks. *Information*, 8(3):88, 2017.

[45] Jie Wu, Hongchun Li, Yi Xu, and Jun Tian. Joint design of wifi mesh network for video surveillance application. In *Proceedings of the 14th ACM International Symposium on QoS and Security for Wireless and Mobile Networks*, pages 140–146, 2018.

[46] Chaojie Yu, Zhipeng Yang, Xiang Chen, and Jian Yang. Scalable video transmission in software defined wireless mesh network. In *2018 4th IEEE Conference on Network Softwarization and Workshops (NetSoft)*, pages 456–461. IEEE, 2018.

[47] Chaofan Yang, Chenshu Wu, Zheng Yang, Tongtong Liu, Zuwei Yin, Yunhao Liu, and Xufei Mao. Enhancing industrial video surveillance over wireless mesh networks. In *2016 25th International Conference on Computer Communication and Networks (ICCCN)*, pages 1–9. IEEE, 2016.

[48] Weixian Liao, Sergio Salinas, Ming Li, Pan Li, and Kenneth A. Loparo. Cascading Failure Attacks in the Power System: A Stochastic Game Perspective. *IEEE Internet of Things Journal*, 4(6):2247–2259, 2017. ISSN 23274662. doi: 10.1109/JIOT.2017.2761353.

[49] V Cagri Gungor, Dilan Sahin, Taskin Kocak, Salih Ergut, Concettina Buccella, Carlo Cecati, and Gerhard P Hancke. A Survey on Smart Grid Potential Applications and Communication Requirements. *IEEE Transactions on Industrial Informatics*, 9(1):28–42, feb 2013. ISSN 1551-3203. doi: 10.1109/TII.2012.2218253. URL http://ieeexplore.ieee.org/document/6298960/.

[50] Juan Pablo Astudillo León and Luis J De la Cruz Llopis. A joint multi-path and multi-channel protocol for traffic routing in smart grid neighborhood area networks. *Sensors*, 18(11):4052, 2018.

[51] Juan Pablo Astudillo León and Luis J. De la Cruz Llopis. Emergency aware congestion control for smart grid neighborhood area networks. *Ad Hoc Networks*, 93:101898, 2019.

[52] Juan Pablo Astudillo León, Thomas Begin, Anthony Busson, and Luis J. de La Cruz Llopis. A fair and distributed congestion control mechanism for smart grid neighborhood area networks. *Ad Hoc Networks*, page 102169, 2020. ISSN 1570-8705. doi: https://doi.org/10.1016/j.adhoc.2020.102169. URL http://www.sciencedirect.com/science/article/pii/S1570870520300974.

[53] Juan Pablo Astudillo León and Luis J. de la Cruz Llopis. Multi channel allocation and congestion control for smart grid neighborhood area networks. In *Proceedings of the 15th ACM International Symposium on Performance Evaluation of Wireless Ad Hoc, Sensor, & Ubiquitous Networks*, PE-WASUN'18, pages 1–8, New York, NY, USA, 2018. ACM. ISBN 978-1-4503-5961-0. doi: 10.1145/3243046.3243050. URL http://doi.acm.org.recursos.biblioteca.upc.edu/10.1145/3243046.3243050.

[54] Juan Pablo Astudillo León, Thomas Begin, Anthony Busson, and Luis J de la Cruz Llopis. Towards a distributed congestion control mechanism for smart grid neighborhood area networks. In *Proceedings of the 16th ACM International Symposium on Performance Evaluation of Wireless Ad Hoc, Sensor, & Ubiquitous Networks*, pages 29–36, 2019.

[55] IEEE Computer Society. *IEEE Std 802.11-2016. IEEE Standard for Information technology—Local and metropolitan area networks. Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications.* New York, 2016. ISBN ISBN 978-1-5044-3645-8.

[56] IEEE Computer Society. *IEEE Std 802.11s, IEEE Standard for Information Technology—Local and metropolitan area networks—Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications. Amendment 10: Mesh Networking.* New York, 2011. ISBN 978-0-7381-6731-2.

[57] Jaebeom Kim, Dabin Kim, Keun-woo Lim, Young-bae Ko, and Sang-youm Lee. Improving the reliability of IEEE 802.11s based wireless mesh networks for smart grid systems. *Journal of Communications and Networks*, 14(6):629–639, dec 2012. ISSN 1229-2370. doi: 10.1109/JCN.2012.00029. URL http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6412861.

[58] Leandro Navarro, Roger Baig Vinas, Christoph Barz, Joseph Bonicioli, Bart Braem, Felix Freitag, and Ivan Vilata-i Balaguer. Advances in wireless community networks with the community-lab testbed. *IEEE Communications Magazine*, 54(7):20–27, jul 2016. ISSN 0163-6804. doi: 10.1109/MCOM.2016.7509374. URL http://ieeexplore.ieee.org/document/7509374/.

[59] Eiman Alotaibi and Biswanath Mukherjee. A survey on routing algorithms for wireless Ad-Hoc and mesh networks. *Computer networks*, 56(2):940–965, 2012. doi: http://dx.doi.org/10.1016/j.comnet.2011.10.011. URL http://www.sciencedirect.com/science/article/pii/S138912861100377X.

[60] Gyanappa A Walikar and Rajashekar C Biradar. A survey on hybrid routing mechanisms in mobile ad hoc networks. *Journal of Network and Computer Applications*,

77:48–63, 2017. doi: http://dx.doi.org/10.1016/j.jnca.2016.10.014. URL http://www.sciencedirect.com/science/article/pii/S1084804516302430.

[61] C Perkins, E Belding-Royer, and S Das. Ad hoc On-Demand Distance Vector (AODV) Routing, RFC 3561 (Experimental). Technical Report 3561, 2003. URL http://www.ietf.org/rfc/rfc3561.txt.

[62] T Clausen and P Jacquet. Optimized Link State Routing Protocol (OLSR). *Internet Engineering Task Force (IETF)*, 4:75, 2003. ISSN 2070-1721. doi: 10.1.1.11.620. URL https://www.ietf.org/rfc/rfc3626.txt.

[63] T. Clausen, Christopher Dearlove, Philippe Jacquet, and Ulrich Herberg. The Optimized Link State Routing Protocol Version 2. Technical Report 7181, apr 2014. URL https://tools.ietf.org/html/rfc7181.

[64] Axel Neumann, Corinna Aichele, Marek Lindner, and Simon Wunderlich. Better approach to mobile ad-hoc networking (BATMAN). *IETF draft, October*, 2008.

[65] BMX6 mesh networking protocol. "http://bmx6.net".

[66] Axel Neumann, Ester Lopez, and Leandro Navarro. An evaluation of BMX6 for community wireless networks. In *2012 IEEE 8th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)*, pages 651–658. IEEE, oct 2012. ISBN 978-1-4673-1430-5. doi: 10.1109/WiMOB.2012.6379145. URL http://ieeexplore.ieee.org/document/6379145/.

[67] a Qayyum, L Viennot, and A Laouiti. Multipoint relaying: An efficient technique for flooding in mobile wireless networks. In *Hal.Inria.Fr*, pages 3866–3875. IEEE, 2000. ISBN 0000000000000. URL https://hal.inria.fr/inria-00072756/document.

[68] Maria Isabel Vara Lorenzo. Descubrimiento de servicios cross-layer basado en OLSR para redes Manet. 2015. URL http://hdl.handle.net/10016/22790.

[69] T. Clausen, C. Dearlove, and J. Dean. IETF RFC 6130, Mobile Ad Hoc Network (MANET) Neighborhood Discovery Protocol (NHDP), IETF, 2011, 2011. ISSN 2070-1721. URL https://tools.ietf.org/html/rfc6130.

[70] C Dearlove, T Clausen, and P Jacquet. Rationale for the Use of Link Metrics in the Optimized Link State Routing Protocol Version 2 (OLSRv2), RFC 7185 (Informational). Technical report, 2014. URL https://tools.ietf.org/html/rfc7185.

[71] C Dearlove and T Clausen. Multi-Topology Extension for the Optimized Link State Routing Protocol Version 2 (OLSRv2), RFC 7722 (Experimental). Technical Report 7722, 2015. URL http://www.ietf.org/rfc/rfc7722.txt.

[72] Ramyar Rashed Mohassel, Alan Fung, Farah Mohammadi, and Kaamran Raahemifar. A survey on Advanced Metering Infrastructure. *International Journal of Electrical Power & Energy Systems*, 63:473–484, 2014. ISSN 01420615. doi: 10.1016/j.ijepes.2014.06.025. URL https://www.sciencedirect.com/science/article/pii/S0142061514003743.

[73] Hamid Gharavi and Bin Hu. Multigate Communication Network for Smart Grid. *Proceedings of the IEEE*, 99(6):1028–1045, jun 2011. ISSN 0018-9219. doi: 10.1109/JPROC.2011.2123851. URL http://ieeexplore.ieee.org/document/5768102/.

[74] Yi Xu and Wenye Wang. Wireless Mesh Network in Smart Grid: Modeling and Analysis for Time Critical Communications. *IEEE Transactions on Wireless Communications*, 12 (7):3360–3371, jul 2013. ISSN 1536-1276. doi: 10.1109/TWC.2013.061713.121545. URL http://ieeexplore.ieee.org/document/6547820/.

[75] Ji-Sun Jung, Keun-Woo Lim, Jae-Beom Kim, Young-Bae Ko, Younghyun Kim, and Sang-Yeom Lee. Improving IEEE 802.11s Wireless Mesh Networks for Reliable Routing in the Smart Grid Infrastructure. In *2011 IEEE International Conference on Communications Workshops (ICC)*, pages 1–5. IEEE, jun 2011. ISBN 978-1-61284-954-6. doi: 10.1109/iccw.2011.5963578. URL http://ieeexplore.ieee.org/document/5963578/.

[76] Jaebeom Kim, Dabin Kim, Keun-Woo Lim, Young-Bae Ko, and Sang-Youm Lee. Improving the reliability of IEEE 802.11s based wireless mesh networks for smart grid systems. *Journal of Communications and Networks*, 14(6):629–639, 2012. ISSN 1229-2370. doi: 10.1109/JCN.2012.00029. URL http://ieeexplore.ieee.org/document/6412861/.

[77] Xiaoheng Deng, Tingting He, Lifang He, Jinsong Gui, and Qionglin Peng. Performance Analysis for IEEE 802.11s Wireless Mesh Network in Smart Grid. *Wireless Personal Communications*, 96(1):1537–1555, 2017. ISSN 1572834X. doi: 10.1007/s11277-017-4255-7.

[78] Yakubu Tsado, Kelum Gamage, Bamidele Adebisi, David Lund, Khaled Rabie, and Augustine Ikpehai. Improving the Reliability of Optimised Link State Routing in a Smart Grid Neighbour Area Network based Wireless Mesh Network Using Multiple Metrics. *Energies*, 10(12):287, feb 2017. ISSN 1996-1073. doi: 10.3390/en10030287. URL http://www.mdpi.com/1996-1073/10/3/287.

[79] Yakubu Tsado, Kelum A A Gamage, David Lund, and Bamidele Adebisi. Performance analysis of variable Smart Grid traffic over ad hoc Wireless Mesh Networks. In *2016 International Conference on Smart Systems and Technologies (SST)*, pages 81–86. IEEE, oct 2016. ISBN 978-1-5090-3718-6. doi: 10.1109/SST.2016.7765637. URL http://ieeexplore.ieee.org/document/7765637/.

[80] Xiaoheng Deng, Lifang He, Qiang Liu, Xu Li, Lin Cai, and Zhigang Chen. EPTR: expected path throughput based routing protocol for wireless mesh network. *Wireless Networks*, 22(3):839–854, 2016. ISSN 15728196. doi: 10.1007/s11276-015-1003-3.

[81] Xiaoheng Deng, Lifang He, Xu Li, Qiang Liu, Lin Cai, and Zhigang Chen. A reliable QoS-aware routing scheme for neighbor area network in smart grid. *Peer-to-Peer Networking and Applications*, 9(4):616–627, jul 2016. ISSN 1936-6442. doi: 10.1007/s12083-015-0331-5. URL http://link.springer.com/10.1007/s12083-015-0331-5.

[82] Xiaoheng Deng, Lifang He, Congxu Zhu, Mianxiong Dong, Kaoru Ota, and Lin Cai. QoS-Aware and Load-Balance Routing for IEEE 802.11s Based Neighborhood Area Network in Smart Grid. *Wireless Personal Communications*, 89(4):1065–1088, 2016. ISSN 1572834X. doi: 10.1007/s11277-016-3305-x.

[83] Xiaoheng Deng, Qionglin Peng, Lifang He, and Tingting He. Interference-aware QoS routing for neighbourhood area network in smart grid. *IET Communications*, 11(5):756–764, 2017. ISSN 1751-8628. doi: 10.1049/iet-com.2016.0860. URL http://digital-library.theiet.org/content/journals/10.1049/iet-com.2016.0860.

[84] Kishwer Abdul Khaliq, Amir Qayyum, and Jurgen Pannek. Performance Analysis of Proposed Congestion Avoiding Protocol for IEEE 802.11s. *International Journal of Advanced Computer Science and Applications*, 8(2):356–369, 2017. ISSN 21565570. doi: 10.14569/IJACSA.2017.080246. URL http://dx.doi.org/10.14569/IJACSA.2017.080246.

[85] Jorjeta Gueorguieva Jetcheva, Sivakumar Kailas, Sachin Kanodia, and Mohan Natarajan. Multi-channel assignment method for multi-radio multi-hop wireless mesh networks, September 2 2014. US Patent 8,824,380.

[86] Sana Ghannay, Sonia Mettali Gammar, Fethi Filali, and Farouk Kamoun. Multi-radio multi-channel routing metrics in ieee 802.11s based wireless mesh networks. *annals of telecommunications - annales des télécommunications*, 67(5):215–226, Jun 2012. ISSN 1958-9395. doi: 10.1007/s12243-011-0253-z. URL https://doi.org/10.1007/s12243-011-0253-z.

[87] ABM Alim Al Islam, Md Jahidul Islam, Novia Nurain, and Vijay Raghunathan. Channel assignment techniques for multi-radio wireless mesh networks: A survey. *IEEE Communications Surveys & Tutorials*, 18(2):988–1017, 2016. doi: 10.1109/COMST.2015.2510164. URL https://ieeexplore.ieee.org/document/7360096.

[88] Mohammad Doraghinejad, Hossein Nezamabadi-pour, and Ali Mahani. Channel assignment in multi-radio wireless mesh networks using an improved gravitational search algorithm. *Journal of Network and Computer Applications*, 38:163 – 171, 2014. ISSN 1084-8045. doi: https://doi.org/10.1016/j.jnca.2013.04.007. URL http://www.sciencedirect.com/science/article/pii/S1084804513001100.

[89] Yeqing Wu, Fei Hu, Sunil Kumar, John D Matyjas, Qingquan Sun, and Yingying Zhu. Apprenticeship learning based spectrum decision in multi-channel wireless mesh networks with multi-beam antennas. *IEEE Transactions on Mobile Computing*, 16(2):314–325, 2017. doi: 10.1109/TMC.2016.2548461. URL http://ieeexplore.ieee.org/document/7444189.

[90] R. Jain, D. M. Chiu, and W. R. Hawe. A quantitative measure of fairnessand discrimination for resource allocation in shared computer system. *Research Report DEC-TR-301, Digital Equipment Corporation, 1984.*

[91] Nenad J Jevtić and Marija Z Malnar. Implementation of etx metric within the aodv protocol in the ns-3 simulator. *Telfor Journal*, 10(1):20–25, 2018.

[92] Junfei Qiu, Qihui Wu, Guoru Ding, Yuhua Xu, and Shuo Feng. A survey of machine learning for big data processing. *EURASIP Journal on Advances in Signal Processing*, 2016(1):67, 2016.

[93] Pratap Chandra Sen, Mahimarnab Hajra, and Mitadru Ghosh. Supervised classification algorithms in machine learning: A survey and review. In *Emerging Technology in Modelling and Graphics*, pages 99–111. Springer, 2020.

[94] Jingjing Wang, Chunxiao Jiang, Haijun Zhang, Yong Ren, Kwang-Cheng Chen, and Lajos Hanzo. Thirty years of machine learning: The road to pareto-optimal wireless networks. *IEEE Communications Surveys & Tutorials*, 2020.

[95] Raouf Boutaba, Mohammad A Salahuddin, Noura Limam, Sara Ayoubi, Nashid Shahriar, Felipe Estrada-Solano, and Oscar M Caicedo. A comprehensive survey on machine learning for networking: evolution, applications and research opportunities. *Journal of Internet Services and Applications*, 9(1):16, 2018.

[96] Wes McKinney. Data structures for statistical computing in python. In Stéfan van der Walt and Jarrod Millman, editors, *Proceedings of the 9th Python in Science Conference*, pages 51 – 56, 2010.

[97] Guido Van Rossum and Fred L Drake Jr. *Python reference manual*. Centrum voor Wiskunde en Informatica Amsterdam, 1995.

[98] Guido Van Rossum and Fred L. Drake. *Python 3 Reference Manual*. CreateSpace, Scotts Valley, CA, 2009. ISBN 1441412697.

[99] Alaa Tharwat. Classification assessment methods. *Applied Computing and Informatics*, 2018.

[100] Mike Papkov. DecisionTreeToCpp. "https://github.com/papkov/DecisionTreeToCpp".