Title	マルチホップ無線ネットワークのためのネットワークコーディングベースの効率的なデータ転送
Author(s)	Nyan, Lin
Citation	
Issue Date	2018-03
Type	Thesis or Dissertation
Text version	ETD
URL	http://hdl.handle.net/10119/15323
Rights	
Description	Supervisor:リム　勇仁，情報科学研究科，博士

Doctoral Dissertation

Efficient Network Coding Based Data Transfer for Multihop Wireless Networks

Nyan Lin

Supervisor: Yuto LIM

School of Information Science
Japan Advanced Institute of Science and Technology

March, 2018

Abstract

With the growing demands of wireless applications and mobile data connections, wireless communication is expected to provide the ever-increasing demand for higher data rate and efficient data communication. Energy consumption is also one of the main problems for the mobile wireless devices because they all depend on the limited battery power installed inside them. Recently, network coding has been introduced as an enabling technology to fulfil the highest achievable capacity of multicast transmissions in butterfly-topology networks. It is also said that network coding has a special advantage to the wireless networks with no extra cost due to the broadcast nature of wireless transmission. This dissertation investigates the potential benefits of network coding to provide efficient data communication for multihop wireless communication and to fulfil the requirements of future wireless networks.

The targeted problems are approached from the viewpoint of multihop communication and the proposed solutions involve the application of network coding techniques. Network coding has a high potential to be incorporated on the Internet in future due to its advantages such as reduction in the number of transmissions and providing high reliability and robustness. However, more work is still needed to realise the practical application of it. This dissertation considers the application of network coding in different scenarios to achieve the efficient data communication.

In this dissertation, an efficient network coding based data transfer framework is proposed, which utilises network coding (NC) to assist data transmission, data collection and data sharing among the wireless nodes to be more efficient in terms of high throughput, low latency, fairness and low energy consumption. Three schemes comprise in the framework for the three scenarios which have high potential to become popular soon. For the first scenario, a network coding-aware medium access control (necoMAC) scheme is developed, which is a combination of network coding-aware 2-hop path selection protocol (NCA-2PSP) and network coding-aware carrier sense multiple access (NCA-CSMA) protocols for the higher rate data transmission from one node to another. These MAC

protocols exploit the multi-rate capability of IEEE 802.11 PHY and golden topologies such as chain and triangle inside the network as key resources for network coded transmissions. These protocols introduce a relay control message, which informs the sender to use a higher rate for transmitting the data frames.

The second scenario includes data gathering applications in a wireless sensor network (WSN), hundreds of nodes are spatially distributed to collect information about the physical environment. Their energy is consumed not only for sensing but also for networking functions to propagate the sensed data to a remote device, base station (BS). The most obvious challenge of a WSN is energy efficiency. Ultra-low latency and ultra-high reliability are also needed for the real-time data collection from the physical environment. For this scenario, an energy-efficient network coding based data gathering scheme called necoDG is proposed. Random linear network coding (RLNC) is applied to the aggregator or cluster head (CH) node to support the reliable data communication from CH to BS. A network coding-aware medium access control protocol is incorporated into cluster-based WSN for the data transmission from each sensor node to the CH node.

For the problem of mobile data downloading from a base station (BS), a balanced cooperative network coded transmission scheme called BCCT is proposed to achieve fast downloading satisfaction. This work includes data sharing within a group of wireless devices by exploiting the two interfaces, cellular and WiFi, which are present in a standard mobile phone. The BCCT scheme applies linear network coding and transmission of the combined data frames to the devices located in the one-hop distance. An algorithm for selecting the next transmitter and choosing the better combination is introduced to maintain fairness among members of the group. The purpose is to satisfy their required data frames as quickly as possible with the fewer number of transmissions.

The proposed schemes are evaluated by computer-based simulation in multihop wireless network environments in terms of the performance metrics such as throughput, energy consumption, latency, fairness and overhead ratio. The results show that the overall performance is improved by 30% compared to conventional methods. The necoDG scheme obtains about 15.2% energy saving on average. The BCCT scheme significantly balances the number of transmissions made from each device and reduces 8.92% of the total number of required transmissions.

Keywords: Network coding, Efficient data communication, Multihop wireless communication, MAC protocol, Cooperative data exchange

Acknowledgments

This work would have not been possible without the support of many people, whom I would like to thank here.

First of all, I am ineffably indebted to my supervisor, Associate Professor Yuto Lim for his invaluable guidance, encouragement and kind help in so many ways. I owe my profound gratitude to my second supervisor, Professor Yasuo Tan for his inspiring instructions and comments throughout this research endeavour. Their brilliant ideas and suggestions helped me to improve this dissertation.

I am extremely thankful to my advisor for the Minor Research Project, Associate Professor Brian Michael Kurkoski for his patient guidance, warm encouragement and useful suggestions to accomplish the Minor Research project and this dissertation.

I would like to convey my sincere gratitude to the members of examination committee: Associate Professor Hiraku Okada from Nagoya University and Associate Professor Razvan Florin Beuran from JAIST, for their patience in reading this dissertation and contributing to its content to improve the quality of the dissertation.

I am thankful to my friends in Tan and Lim laboratories for the time we spent together and gatherings throughout the years and for the valuable discussions on various topics.

I am extremely thankful to Government of Japan, JAIST and Japanese people for their financial support by the MEXT scholarship programme offered to my country. I am also immensely grateful to Ministry of Education and Myanmar Government for their permission to study in Japan.

Finally, I would like to appreciate my family, relatives and friends from Myanmar for their continuous love and encouragement.

Contents

List of Figures

List of Tables

Acronyms

2PSP	2-hop Path Selection Protocol
AODV	Ad hoc On-Demand Distance Vector
AWGN	Additive White Gaussian Noise
BCCT	Balanced Cooperative Coding and Transmission
BER	Bit error rate
AP	Access Point
BS	Base Station
CH	Cluster Head
CSMA/CA	Carrier Sense Multiple Access with Collision Avoidance
CTS	Clear-To-Send
CW	Contention Window
DCF	Distributed Coordination Function
DIFS	Distributed Inter Frame Spacing
D2D	Device-to-Device Communication
IoT	Internet of Things
LEACH	Low-energy Adaptive Clustering Hierarchy
M2M	Machine-to-Machine Communication
MAC	Medium Access Control
MSDU	MAC Service Data Unit
MWN	Multihop Wireless Network
NAV	Network Allocation Vector
NCA-CSMA	Network Coding-aware Carrier Sense Multiple Access
NCA-2PSP	Network Coding-aware 2-hop Path Selection Protocol
NDBPS	Number of Data Bits per OFDM Symbol
necoMAC	Network Coding-based Medium Access Control
necoDG	Network Coding-based Data Gathering

Acronyms

OFDM	Orthogonal Frequency Division Multiplexing
OSI	Open Systems Interconnection
PLNC	Physical Layer Network Coding
QoS	Quality of Service
RCTS	Relay-Request-To-Send
RLNC	Random Linear Network Coding
RRTS	Relay-Request-To-Send
RTR	Request-To-Relay
RTS	Request-To-Send
SIFS	Short Inter Frame Spacing
SINR	Signal-to-interference-plus-noise ratio
TMT	Theoretical maximum throughput
WSN	Wireless Sensor Network
XOR	Bitwise Exclusive OR

List of Symbols

Symbols in Chapter 5

c_i	i^{th} client device
x_i	i^{th} packet
X	The set of incoming packets
C_x	The set of linear combinations of packets
Y_i	The subspace spanned by vectors corresponding to packets
b	A vector
P_{init}	Initial packets receiving probability
GF	The finite field elements

Chapter 1

Introduction

This chapter describes the overview of the dissertation. Firstly, Section 1.1 explains the growing demand for high-performance wireless communication in the near future. Based on the requirements of future wireless networks, some of the problems related to energy consumption, latency and number of transmissions are to be solved in this dissertation. These targeted problems and the focus of research are presented in Section 1.2. The approaches to the selected problems and the objectives of the dissertation are described in Section 1.3 and Section 1.4, respectively. The main contents of each chapter are briefly described in Section 1.5 Dissertation organisation.

1.1 The Growing Demand for High-Performance Wireless Communication

Information and communication technologies (ICT) have contributed to the social and economic development of many countries these days and worlds business model is transforming to digital. A digital economy is rapidly replacing the whole worlds economic system and it has strong impact on the economic and social development. In addition, the social development leads to the changes in the way mobile and wireless communication systems are used. The growth in global consumer mobile services such as mobile banking and commerce, mobile social networking, mobile gaming, mobile email, mobile music and video will surprisingly increase by 2021 according to the estimation of Cisco® Visual Networking Index (VNI) [1]. The traffic from wireless and mobile devices will

become 63% of Internet traffic with only 37% from wired devices.

The expectation of future wireless applications includes very high demand of high data rate, high availability, and low latency for applications such as media streaming applications and healthcare systems. The coexistence of human-centric and machine-type applications leads to diverse requirements on mobile and communication systems. More stringent latency and reliability requirements are needed for the applications related to healthcare, mission-critical control systems and security. The scalability and flexibility are also important due to the existence of many connected devices with diverse applications. The challenge is the trade-off between satisfying the requirements and the growing cost. Efficiency and scalability become the key design criteria [2].

It is estimated that the smartphone traffic will exceed the PC traffic with 33% of total IP traffic in 2021 although it reached only 13% in 2016. The traffic from PC will decrease to 25% of total IP traffic in 2021. There will also be an increased number of devices connected to IP networks in 2021 and it will become three times of the global population at that time. It means that a person will have 3.5 networked devices and the IP traffic per capita will reach 35 GB per capita. Among them, machine-to-machine (M2M) applications such as healthcare monitoring and smart meters will also have an increase of 2.4 times that of the current number and will become 13.7 billion connections by 2021.

The current trend and challenges of future generation wireless communication are motivating the researchers and the industry for the new revolution Fifth Generation mobile technology, 5G. The 5G is viewed as providing communication and data services using all possible access solutions and core network switching rather than a new radio access technology. Cooperative communications and network coding, full duplex, massive multiple inputs and multiple outputs (MIMO), device-to-device (D2D) communications and green communications are some of the promising techniques for 5G [3].

1.2 Problems and Research Focus

1.2.1 Energy-Efficient High-Rate Communication

With the growing demands of wireless applications and mobile data connections, wireless communication is expected to provide the ever-increasing demand for higher data rate and efficient data communication. On the other hand, the total energy consumed for wireless communication systems is more than 3% of the worldwide electric energy consumption and it is expected to increase rapidly in the future [4]. Mobile terminals in wireless systems also have energy limitation. Although the constraints on computation and storage can disappear with the development of fabrication techniques, energy limitation will still be challenging [5]. The battery technology does not improve fast enough to catch up the development of information and communication technology. Therefore, achieving high data rate with low energy consumption (Mbps/Joule) becomes one of the main problems for the mobile wireless communication.

1.2.2 Low Latency and High Reliable Communication

With the advancements in the Internet of Things (IoT), computation and communication technologies, the world is leading to build the connected devices, not limited inside the industries but smart streets, smart supermarkets, smart homes and even smart cities too. It has the advantages in many areas such as agriculture, health care, production, better services by analysing and monitoring the feedback data from those devices, and leads to more economical world in a way of saving the resources. In this scenario, massive deployment of sensors and actuators is a typical application. Connectivity and longer battery life are requirements for the collection of updated data from the physical environment. The traditional way of reliable communication is based on the retransmission of packets and acknowledgements. It costs bandwidth and high latency for high number of retransmissions to achieve the successful data transfer. As a requirement for future application scenarios, ultra-low latency, and ultra-high reliability are needed.

The available communication technologies such as Wi-Fi, Bluetooth, ZigBee and 2G/ 3G/ 4G cellular depend on the application and factors such as range, data requirements, security, power demands and battery life [6]. Wi-Fi connectivity is often an obvious choice

because there is wide existing infrastructure as well as offering fast data transfer. The approximate range of connectivity is 50 meters. For the operation of longer distances, the IoT applications can take advantage of GSM/ 3G/ 4G cellular communication capabilities. However, these technologies cost very high power consumption for many IoT applications. For wide area network (WAN) applications, another choice is LPWA technologies [7] which can provide low power consumption and support large networks with millions of devices in low data rates range from 0.3 kbps to 50 kbps.

1.2.3 Quick Download Access in a Fewer Number of Transmissions

Another scenario which has a high prospect for future application is that end-user needs connectivity even in very crowded places such as stadiums, shopping malls, and public events. Due to the increasing number of connected devices to the IP network and high data rates services such as video streaming and file downloads, the traffic volume is very high and leads to network overload. As a result, users suffer from service denials. Assume that some nearby wireless client devices download a file of n packets or some gaming from a common base station. Some devices may receive the whole file while some may receive only a portion of it and lose other subset of the file. If they separately request the lost packets to the base station, the number of required retransmissions will be high and it will cause more burden on the base station in large crowds. In this case, the nearby wireless devices can utilise the advantage of short-distance communication links by creating a cooperative data sharing group among them via the built-in Bluetooth or Wi-Fi interfaces. This scenario has been considered in [8–11].

With this approach, devices will possess a faster and reliable short-range communication service to provide their requirements without any request to a far base station again. The network overload can also be reduced and base station will be able to serve other users. The clients only need to collectively possess n packets within their group and their desire to cooperate to share the packets they have already received. This concept is very similar to peer-to-peer (P2P) networking in the viewpoint that the resources such as processing power, storage or network bandwidth are made available to other participants without the need for central coordination by servers.

These scenarios are illustrated in Figure 1.1. The figure is originally from the mobile and wireless communications enablers for the twenty-twenty information society (METIS) project [2]. METIS also believes wireless network coding and buffer-aided relaying to be the promising research directions in multi-node/multi-antenna transmission.

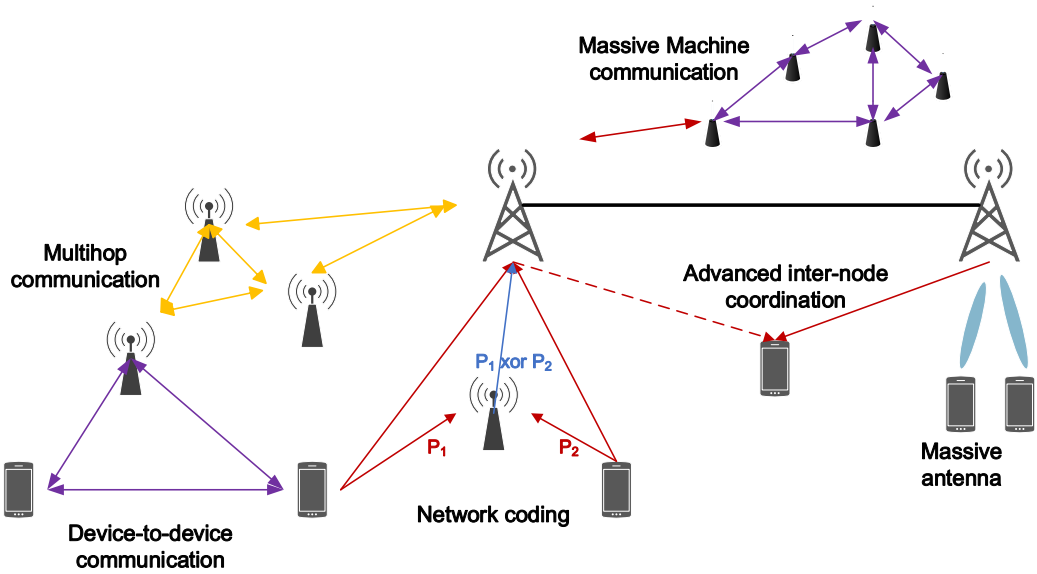


Figure 1.1: Multi-node/Multihop Communication

All the scenarios include multi-hopping as a common communication paradigm. This paradigm is currently not a main option in IEEE 802.11 standard due to its remaining issues such as high latency and its effect to throughput improvement. However, there is a high chance to become a core communication paradigm in future because of the dense usage of wireless devices and the new techniques such as wireless network coding. It can be setup in ad hoc fashion or via infrastructure such as cellular stations. Multihop relay-based communication is gaining global acceptance as one of the most promising technologies in next-generation wireless cellular networks with the performance expectations such as throughput improvement, an extension of coverage area and a decrease of energy consumption [12, 13]. Network coding techniques can be integrated into the relay-based multihop wireless networks to reduce the remaining weaknesses and can benefit the remaining issues of multihop wireless communication. For the above reasons, this dissertation specifically focuses on the multi-node and multihop communication among

the very diverse technologies to achieve the targeted performance and capability of future wireless systems.

1.2.4 Multihop Wireless Communication and Research Issues

Multihop communication has a strong research background in the relay-based transmission [14]. However, there are some remaining issues. Due to operation in half-duplex manner, there is inefficiency in spectrum usage. Multiple time slots are required and consequently, it affects the throughput improvement. Another problem introduced in multihop communications is the latency. Relay-based communication also consumes communication resources, as more hops are needed. More researches are needed to find out more solutions, which are efficient, reliable, less energy consumed, and able to provide the demand of todays wireless applications.

Although each wireless device is equipped with cellular or Wi-Fi interfaces, it also possesses Bluetooth, which is not as frequently used as cellular or Wi-Fi interfaces for communication purpose except for sharing files between two nearby devices. Nowadays, nearly everybody has a mobile phone in his/her pocket while commuting outside or working inside the building. Therefore, there is a high chance of connectivity for communication among these existing dense wireless devices.

Emergence of wireless network coding (NC) techniques has brought a new life to the relay-based solutions. Wireless network coding allows the intermediate nodes to combine the packets for multiple independent communication flows with no extra cost because of the broadcast nature of wireless transmission. The advantage of network coding is the reduction in the number of transmissions without affecting the recovery of original messages at the destination nodes. Network coding has high potential to be integrated into future wireless networks to increase throughput and to save energy and bandwidth resources.

1.2.5 Network Coding and Its Potential to Future Wireless Networks

Network coding (NC) is a technique of treating the information packets in network as combinations of each other by some appropriate methods rather than the simple com-

modity data flow. From this fundamental concept, the field of network coding was born in the year 2000 from a seminal paper about network information flow in [15].

Since that time, research on network coding has become a new area of study and many researchers revealed the potential of network coding for the future wireless networks. It can be said that network coding is a portion of routing function at the network layer as it is a concept of combining incoming packets at routers before they are forwarded to the next hop. Recently, network coding has been introduced to apply at the data link layer to obtain more advantages from it and to support the upper layers. The broadcast nature of the wireless transmission benefits the overhearing of the messages at the neighbour nodes and these overheard messages can be efficiently utilised as a coding opportunity for network coding. Network coding has the potential for the future wireless communication because of its special advantage to the wireless networks with no extra cost for broadcast transmission.

The recent researches focus on the implementation challenges and implications of network coding for the practical applications in current and future wireless networks in order to exploit the rich properties of network coding technique such as reduction of required number of transmissions, improvement in throughput, improving the reliability for lossy networks, and saving energy and bandwidth. Depending on the nature of applications and network environment, several different ways of network coding techniques and schemes are developed, and even incorporated into the existing technologies. COPE, the first practical work for the networks with perfect links [16], I^2NC, a joint inter- and intra-session network coding scheme for wireless networks with lossy links [17], and MORE, a wireless opportunistic routing protocol with network coding [18] reveal the promise of network coding for wireless communication.

Moreover, IRTF creates a Network Coding Research Group called NWCRG [19] with the objective to research network coding principles and methods that can benefit the Internet communication. The goals of NWCRG are to gather information on the existing practical implementations of network coding and to propose a path to standardisation of network coding-enabled communication. Therefore, network coding has a bright future with the research and market potential for future communication networks.

In combination with the high potential of D2D communication, network coding has

become an attractive solution to harness the power of wireless and cooperative networks in order to provide communication efficiency, higher throughput and lower energy expenditure. D2D communications has high potential to be integrated to the 5G technologies due to the growing number of devices to be connected in the future. D2D communication creates a market potential for new services and new approaches as it can provide end user benefits such as reduction of power consumption, increase in throughput, and operator benefits such as spectrum efficiency and extension of coverage [20]. Due to the advantage of dense network of wireless devices and popularity of various wireless applications, multihop relaying with the help of network coding will become the key way of communication for future again.

Although network coding is not widely applied in practice yet due to its limitations such as complexity for computing capabilities and security vulnerabilities, some of the real-life measurements have been demonstrated on the commercial devices such as Xperia and S3 [21]. It shows that the energy used for coding and decoding is much smaller than that used for transmission and reception. Moreover, one of the key telecom players, Nokia, displayed the network coding technology at their booth in Mobile World Congress (MWC'17) for using it in the Nokia 5G network. The innovations with network coding are leading to integrate with 5G. For example, a recent article describes how ultra-fast 5G will enable to control the V2V networks and IoT applications by using RLNC technique [22].

Communication protocols should not be built in the usual way by looking at a single communication flow at a time, but rather multiple communication flows should be processed jointly. The new challenge is the proper combination of network coding techniques with multihop relay concept for the scenarios discussed in the requirements of future generation networks. In this dissertation, an investigation is made on how the benefits of network coding can be utilised in multihop wireless networks to provide efficient data communication and to fulfil the requirements of future generation networks.

1.3 Motivation and Vision

This dissertation investigates the potential benefits of network coding to provide efficient data communication for data transmission, data collection, and data sharing in future

wireless networks, especially in multihop wireless communication. The intention is to find the research-based knowledge for the possibility of practical implementation in future wireless technologies.

The vision is to contribute an efficient and fair future wireless communication by using network coding techniques.

The research issue relating to network coding and medium access control is due to the fairness of DCF mechanism in IEEE 802.11 standard [23]. To achieve the highest benefit from network- coded packets/frames, they need to be transmitted with a higher priority than the normal packets/frames. This issue becomes very clear in [16], which proposes the first practical inter-session network coding approach called COPE for the networks with perfect links. The relay node transmits the XOR-coded packets and reduces one transmission time slot like in the proposed NCA-CSMA protocol. As COPE relies on the DCF operation of IEEE 802.11, relay node needs to compete the channel access with other nodes to transmit the coded packet. Therefore, performance of network coding opportunity is totally depending on DCF and is limited by DCF.

The application of network coding into layer 2 of OSI reference model appears in [24]. The author shows that Layer 2.5 NC works best in high-rate symmetric traffic flows between two clients connected to the same 802.11 access point. It can be observed that symmetric traffic flows and high rate are essential to gain benefits from using networking coding at AP. This observation is utilised in the proposed NCA-2PSP protocol, where network coding functionality is applied at the relay node identified by the 2PSP mechanism. A network coding-aware medium access control (necoMAC) scheme is also proposed by combining some MAC protocols with NC.

Relating to the data gathering application, there is an important issue relating to the energy efficiency although many data aggregation techniques and routing protocols have been designed for a number of performance metrics such as energy efficiency, reliability, quality of service. It is the error and loss in wireless communication due to the dynamic channel conditions.

The traditional way to provide reliability is to use the feedback messages to report the received or lost packets. However, these feedback messages consume bandwidth and energy. To prolong network life in resource-limited network and wireless communication,

energy-efficient error control techniques remains a challenge. It is important to provide reliable communication to improve network lifetime of sensor networks.

To solve this problem, there are some design choices with network coding technique. Network coding opportunities may exist within the clusters or outside the clusters along the path to the Base station (BS). For example, the gateway can perform as data aggregator, using some network coding technique to reduce the transmission, or some nodes in the vicinity can be nominated as an aggregator or network coder to provide short-range communication locally. This way can recover loss on long-distance wireless link and can improve high reliability. The cluster heads become the relays that decide the network coding transmission among the other cluster heads and the client nodes. An energy-efficient network coding based data gathering scheme called necoDG is proposed.

The scenario of cooperative data sharing has been considered in [8–10]. The goal is to fulfil the requirements of all participants in the group with minimum transmissions. The encoding scheme that will minimise the number of transmissions is referred to as index coding [25], where a central base station performs the transmissions of linear combined packets to other clients. With linear network coding [26], benefits of cooperation can be further considered for the data exchange problem because many devices can simultaneously gain from one linearly coded packet (linear combination) transmission. The goal is to increase the required packets of as many nodes as possible with the minimum number of transmissions.

In the previous works, a candidate is selected as a transmitter for next round based on the maximum number of received packets. If there is more than one candidate, those schemes choose the next transmitter randomly. In this dissertation, fairness on the number of transmissions is considered in each client while also maintaining the minimum number of transmissions as a whole. A balance coded transmission scheme called BCCT is proposed by exploiting the random linear network coding technique. The reason is to balance the responsibility of transmission from each client and to maintain energy saving for each individual participant. Like in the case of necoDG, the responsibility as a cluster head is rotated among the members of a cluster to avoid the energy depletion of a certain device. If a particular device with an independent packet runs out of energy quickly, other devices will not satisfy their needs and the data exchange process cannot accomplish.

The approach to the three problems mentioned above includes designing a required protocol and advancing the required algorithms for incorporating network coding techniques to work properly with the existing characteristics of the networks for each data transferring scenario.

This approach consists of designing new MAC protocols, algorithms and schemes that can exploit the benefits of network coding to achieve performance metrics such as high throughput, low latency, fairness and low energy consumption. A data transfer framework called E-neco is proposed, it includes three schemes for each of the problem discussed above. They are network coding-aware medium access control (necoMAC) scheme, network coding based data gathering (necoDG) scheme for WSNs, and balanced cooperative coding and transmission with physical layer network coding (BCCT/PLNC) scheme for high rate data exchange with low power consumption between a cellular base station and its mobile stations. It is expected that the three different data communication processes can be performed in a way to achieve higher efficiency, reduce energy and bandwidth resources. This framework is intended to provide the data transferring services such as data transmission, data collection and data sharing in a more efficient way in terms of key performance indicators such as energy and bandwidth efficiency, reliable and low latency communication, and low energy consumption.

It also finds a way of applying the main methodologies at the MAC layer such as:

1. Relay based higher-quality path selection mechanism such as 2PSP with overhearing
2. Adaptive coding for different topologies such as golden chain and golden triangle
3. Technologies based on relaying, multi-hopping and wireless network coding such as RLNC and XOR to achieve the goals in a simpler and flexible way

Preliminary assumption for this framework is that the wireless interference between the nearby nodes is cleared by some sort of interference management techniques. Among the very diverse technologies to achieve the performance and capability targets of 5G wireless systems, this research specifically focuses on the multi-node and multihop networks as shown in Figure 1.1.

Many benefits can be expected from this work. They are throughput improvement, transmission utilising multi-rate/multihop, fairness, reducing number of transmissions,

keeping energy saving and providing long battery life, user benefits with quick download completion, reducing the burden on base station, low latency, and reliable communication.

1.4 Purpose and Objectives

The main purpose of this dissertation is to investigate the potential of network coding technique in a multihop wireless network environment focusing mainly on the data link layer transmission. To achieve the throughput improvement, fairness among the participants, a fewer number of transmissions, low latency and energy efficiency for the data communication in a network, a network coding based data transfer framework is proposed. By proposing a MAC protocol for data transmission between two nodes, data gathering process in a wireless sensor network (WSN) and group-based data sharing scenario between a base station and a group of wireless nodes. This work is intended to explore more scientific knowledge of network coding and its potential applications for the future wireless communication to help the designers and developers of protocols and products, applying the findings of this work. To state clearly and concretely, the purpose and objectives of this dissertation are defined as follows.

Purpose: To incorporate network coding technique in Multihop Wireless Communications for the efficient data transfer.

Objectives:

- To propose a framework for efficient network coding based data transfer between the MAC layer and network layer of MWNs. (Material related to this objective appears in as yet unpublished paper [1] and published paper [10].)
- To introduce a network coding-aware MAC (necoMAC) scheme for efficient data transmission in MWNs. (Material related to this objective appears in published papers [5], [8] and in yet unpublished paper [3].)
- To present a network coding based data gathering (necoDG) scheme for efficient data collection in WSNs. (Material related to this objective appears in published papers [4] and [7].)

- To establish a Balanced Cooperative Coding and Transmission (BCCT) Scheme for efficient data sharing in cellular networks. (Material related to this objective appears in published papers [6], [9] and in yet unpublished paper [2].)

1.5 Dissertation Organisation

The remainder of this dissertation is organised as follows.

- Chapter 2. E-neco Framework
  Chapter 2 begins with a background of network coding and multihop wireless networks. The classification of network coding techniques and schemes are presented based on the existing work in the literature. Different types of topologies are also discussed relating to the network coding opportunity. The medium access control protocols and the overview of the efficient network coding-based data transfer (E-neco) framework are explained in the later part of this chapter.

- Chapter 3. Network Coding-Aware MAC (necoMAC) Scheme
  The design and implementation of network coding-aware medium access control (necoMAC) scheme is described in this chapter. This necoMAC scheme is intended for the data transmission in multirate multihop wireless networks. Design of network coding-aware 2-hop path selection protocol (NCA-2PSP) and network coding-aware carrier sense multiple access (NCA-CSMA) protocols are described.

- Chapter 4. Network Coding Based Data Gathering (necoDG) Scheme
  Chapter 4 describes an energy efficient data gathering scheme with network coding. The random linear network coding (RLNC) is integrated at the cluster head (CH) or the aggregator of each cluster to assist reliable data transfer to the base station (BS). The energy-efficient 2-hop path selection protocol is also introduced for data transmission from sensor nodes to the CH in a cluster.

- Chapter 5. Balanced Cooperative Coding and Transmission (BCCT) Scheme
  A network coding-based cooperative data sharing scheme is implemented in Chapter 5. A group of nearby wireless nodes helps each other to provide the needed packets. This research includes the application of the combination of linear and physical layer

network coding. The goal is to increase the required packets of as many nodes as possible with the minimum number of total transmissions. The BCCT scheme is to maintain the fairness on the number of transmissions made from each device and to reduce the total number of required transmissions. The scheme utilises overheard information so that each client can choose a combination that will be independent from the previously transmitted combinations.

- Chapter 6. Conclusion and Future Work

  Chapter 6 summarises the findings, concludes the dissertation and draws some future trends.

Chapter 2

E-neco Framework

2.1 Introduction

This chapter presents some background on network coding and multihop wireless networks. The classification of network coding techniques and schemes are also presented based on the existing work in the literature.

The overview of the proposed efficient network coding-based data transfer (E-neco) framework is also explained in the later part of this chapter.

2.2 Network Coding and Its Applications

Network coding is a technique which can be applied at a source or at an intermediate node in the network to create new outgoing packets by processing the data packets by some mathematical functions. It can replace the traditional 'store and forward' paradigm at an intermediate node. It was born in 2000 in a seminal paper by Ahlswede et al. [15] in which intermediate nodes or routers were viewed as locations which could provide better throughput by additional computational tasks. The main idea behind this concept is that instead of considering the node as a simple information flow relay, network coding allows the node to combine several packets into a single packet.

Network coding was initially a purely theoretical approach and earlier research is more focused on understanding fundamental theoretical properties and its limits. There was significant progress on this front and as a result, the computational complexity and

efficient network coding algorithms have been discovered. Much work in network coding has concentrated on Random Linear Network Coding (RLNC) technique [27–29] and its application to wireless networks for the improvement of network performance and reliability. The basic idea is that intermediate nodes form linear combinations of the received packets and send the linearly combined packets to the next hop node. In order to extract the original messages at the receiver, the decoding process needs to have a sufficient number of independent linear combinations. This technique requires an extra overhead of carrying coding coefficients in the packet header and computational complexity when larger finite fields (Galois field, GF) are used for the coding coefficients. Its counterpart, XOR coding is relatively simple for coding and decoding processes. Network coding has many potential benefits such as an increase in throughput and an improvement in the reliability and robustness of the network [30]. Many researches on network coding have been carried out and their results have revealed the potential benefits of network coding as well.

A good introduction to the theory and application of network coding is explained in a paper [26] and in a tutorial from the perspective of network coding as a generalisation of routing [31]. A survey and classification of the coding algorithms for wireless multicast transmission is systematically presented in [32] and classification of network coding and its applications are explained with illustrative examples in the survey paper of [33].

2.2.1 Classification of Network Coding Techniques

Since network coding has potential benefits of energy and bandwidth savings for the communication networks, the research community has investigated many approaches to improve the performance of IP networks and cellular networks by incorporating network coding into existing technologies. Depending on the different approaches for various purposes, network coding techniques can be classified into network layer approaches, link layer approaches and the combined application of core network coding techniques and schemes.

2.2.1.1 Incorporation of Network Coding into Different Layers

- ***Network Layer Network Coding***

  Network coding is firstly proposed as a branch of the routing protocol in the network

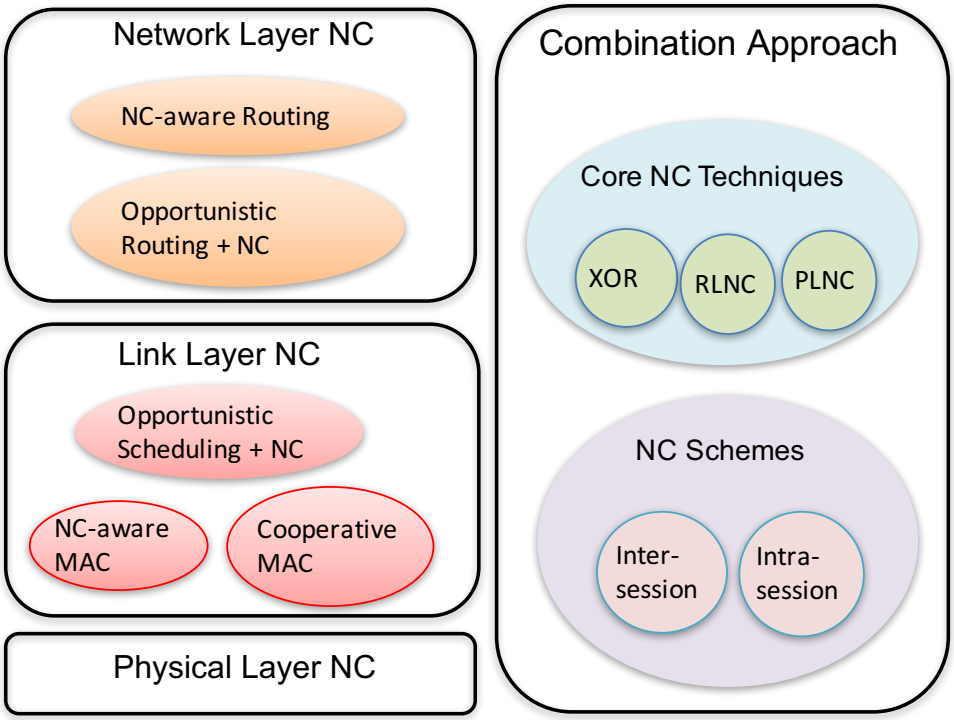


Figure 2.1: Classification of Network Coding Techniques

layer, where traditional store-and-forward communication fails to achieve the multicast rate in a butterfly network. Much research focuses on the network coding-aware routing protocols, in which network coding opportunity is considered as a metric in order to exploit the benefits of network coding. A distributed coding-aware routing protocol DCAR [34] is a kind of network coding-aware routing. In opportunistic routing, the next-hop node is chosen on-the-fly instead of assuming a complete path determined in advance by a routing protocol. The number of control messages required to find next-hop node is high. Network Coding is applied in opportunistic networks to reduce the number of transmissions in cooperation between neighbour nodes to find a path. Network Coding-aware routing protocols are improving the end-to-end delay, network throughput, and reliability of the network [35].

- ***Link Layer Network Coding***

  Researchers found that network coding can also be applied in the link layer so that the benefits of network can be achieved from each link-level transmission and improves the performance of the overall network. Most of the research concerns with the network coding-aware medium access control techniques, opportunistic

scheduling with network coding and cooperative medium access control [36, 37]. To exploit the benefits from a network coder node, the scheduling and medium access is important.

2.2.1.2 Core Network Coding Techniques

- ***XOR***

  A simple network coding technique that uses bit-wise addition over the entire length of the frame, and is one of the most easily understood network coding schemes. When a node receives packets for two different sources, it simply XOR them together and send the combination to outgoing links. A receiving node must maintain a record of at least $n-1$ original packets to be able to decode the received combination. The decoding is simply performing XOR function of the received packet with its stored packet. An example XOR operation in a butterfly network topology is shown in Figure 2.2(a), where a sender 'S' multicasts two messages A and B to two destination nodes 'D_1' and 'D_2'. Each link can carry a data message at a time. As there are two incoming messages at the middle intermediate node 'I_1', a bottleneck at the outgoing link prevents two destinations from receiving both messages A and B simultaneously. This problem can be solved by applying XOR network coding on the two messages and transmitting the coded message on the outgoing link.

- ***Random Linear Network Coding (RLNC)***

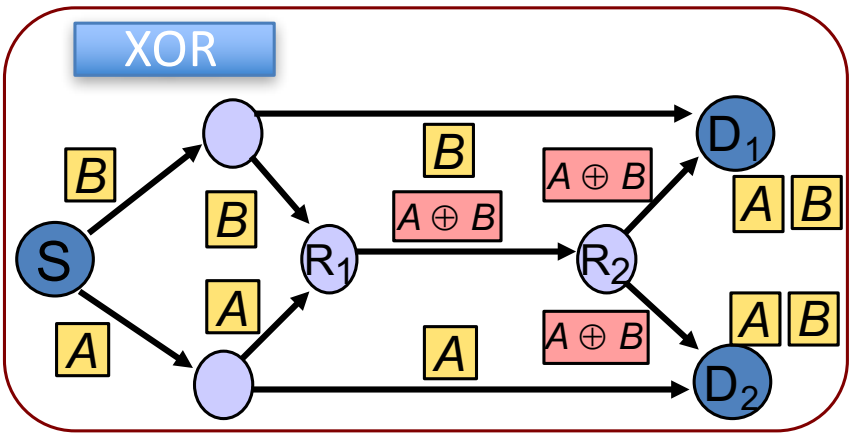
  RLNC [29] is a well-known intra-session network coding technique. In RLNC, when a node receives a packet, it stores the packet in its memory. When packet forwarding occurs at an outgoing link of the node, the node creates the encoded packet from a random linear combination of the packets in its memory. Suppose the node has L packets $u_1, u_2, \ldots, u_L$ in its memory. Then the encoded packet formed is

$$u_0(t) = \sum_{i=1}^{L} \alpha_i u_i \tag{2.1}$$

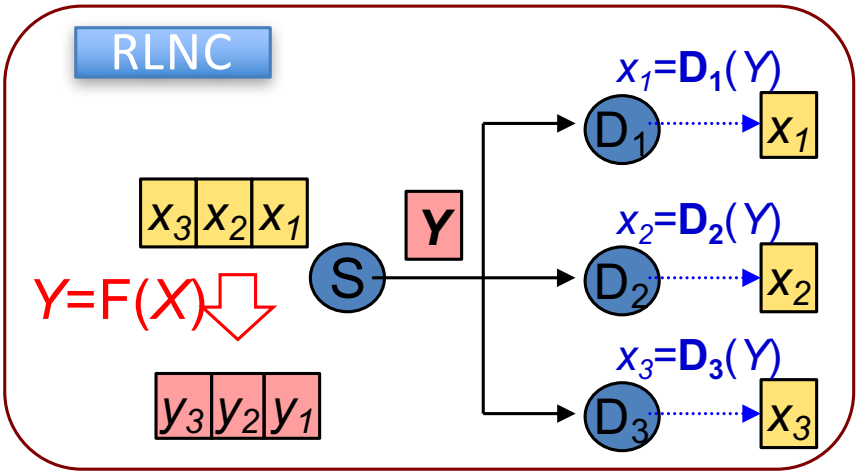
  where i is chosen according to a uniform distribution over the elements of some finite field F_q. The packet's global encoding vector γ is placed in its header and which satisfies

$$u_0(t) = \sum_{k=1}^{K} \gamma_k w_k \tag{2.2}$$

where the w_k are the message packets from the source. For the decoding process, the sink node performs Gaussian elimination on the set of global encoding vectors from the packets in its memory to find the original messages. Figure 2.2(b) shows an example operation of RLNC. A sender, a source or an intermediate node, indicated by 'S' in the figure performs a random linear network coding function, $F(x)$ over the packets x_1, x_2 and x_3 to create an encoded packet Y, which includes some symbols y_1, y_2 and y_3. Then, Y is transmitted to the destination nodes D_1, D_2 and D_3 where the original packets x_1, x_2 and x_3 are recovered by Gaussian elimination operation on Y.



(a)



(b)

Figure 2.2: Types of network coding (a) XOR in a butterfly network (b) Example of RLNC operation

- ***Physical Layer Network Coding (PLNC)***

  The basic idea of PLNC is to exploit the mixing of signals that occurs naturally when electromagnetic (EM) waves are superimposed on one another [38]. The simultaneous transmissions by several transmitters result in the reception of a weighted sum of the signals at a receiver. Their addition will be relayed by an intermediate node. This weighted sum is a form of network coding operation by itself.

2.2.1.3 Network Coding Schemes

Network coding can be classified as either inter-session or intra-session depending on the way of encoding the incoming packets. Inter-session network coding allows the packets from different sessions (sources) to be mixed before forwarding into the next hop. In contrast, intra-session network coding can be used to mix the packets from the same session to address the packet loss problem.

- ***Inter-session Network Coding Scheme***

  If the intermediate wireless nodes are allowed to code the incoming packets, the broadcast nature of wireless nodes becomes an opportunity. Inter-session network coding solves the bottleneck problem and reduces the number of transmissions, by allowing packets from different sessions (sources) to be coded together. By reducing the number of required transmissions, network coding increases the throughput and decreases the interference between the links in wireless networks [39].

- ***Intra-session Network Coding Scheme***

  Another important application of network coding is to provide reliability in wireless networks. The traditional way to provide reliability for both wired and wireless networks is to use feedback messages to report the received (or lost) packets. Network coding can provide reliability with a fewer transmissions by coding the packets from the same session (source). This type of coding is called intra-session network coding.

2.2.2 Topology for Network Coding Opportunity

Topology and group formation mechanism play important roles for the improvement of system performance. One of the three main components of E-neco framework is topology

component. This framework focuses on the utilisation of network coding at the link-level data transfer operations. There are limitations of network coding opportunity at network layer to achieve the maximum performance because the actual transmissions can only happen when the medium is occupied. Therefore, medium access control influences the network coded transmission.

Topology influences the performance of network coding regarding several important metrics. It also affects the performance of control algorithms for scheduling of transmissions, routing, and broadcasting in an ad hoc network [40]. The network coding opportunity also depends on topology where the source, relay and destination form a specific structure in wireless networks. For example, relay node performs XOR network coding to combine some packets in its incoming buffer based on the information of its neighbours.

There are certain types of structures that will benefit the network coding opportunity. These include golden chain (two-way relay channel), golden triangle (relay channel), cross (Star), butterfly and Y (multiple access relay channel) topologies from network perspective. Another type of useful topology for group formation and connections among group members is mesh topology.

Analysis of the relationship between topology and network coding for each type is briefly presented in this section.

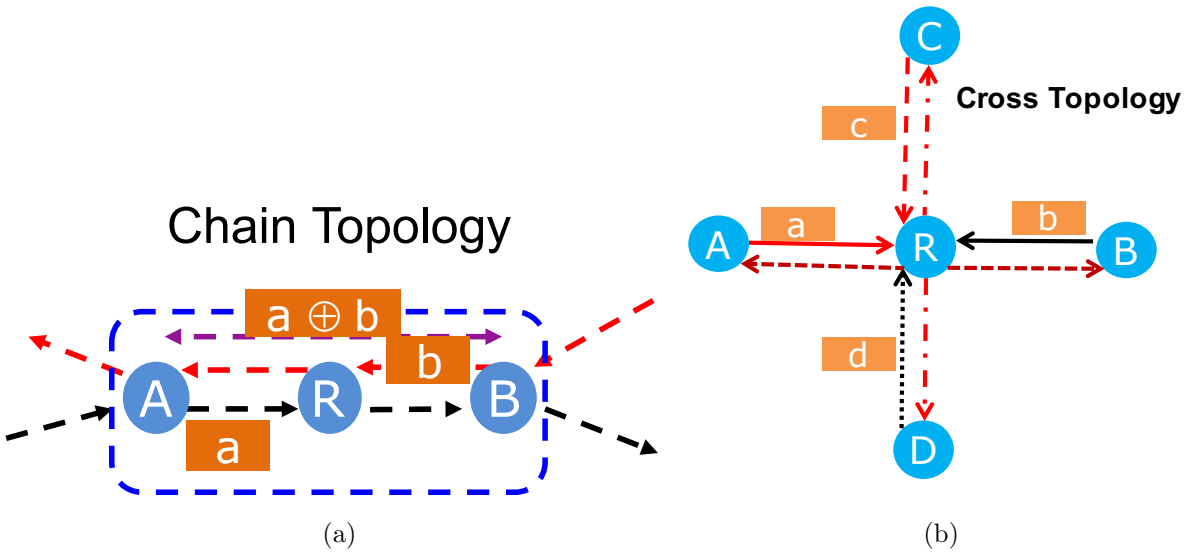


Figure 2.3: Chain topology (two-way relay channel) (b) Cross (star) topology

In Figure 2.3(a) and Figure 2.3(b), chain and cross topologies are illustrated. A chain

topology includes three nodes located in a way that the intermediate node 'R' can hear transmissions from both nodes 'A' and 'B', but 'A' and 'B' cannot hear each other. XOR network coding can be used at 'R' to reduce the number of transmissions in the message exchange between 'A' and 'B'. A cross topology can be seen as two chains cross each other at the common intermediate node.

In a triangle topology (Figure 2.4(a)), the positions of the three nodes 'A', 'R' and 'B' are in a triangle shape. The advantage of this topology is that both the intermediate node 'R' and the receiver 'B' can hear from sender 'A'. Transmissions in high rates can be achieved via the relay 'R' for the throughput improvement of transmission of payload data. In Figure 2.4(b), a multiple access relay channel is depicted. This looks like two triangles with a common relay to forward the incoming data to the same destination.

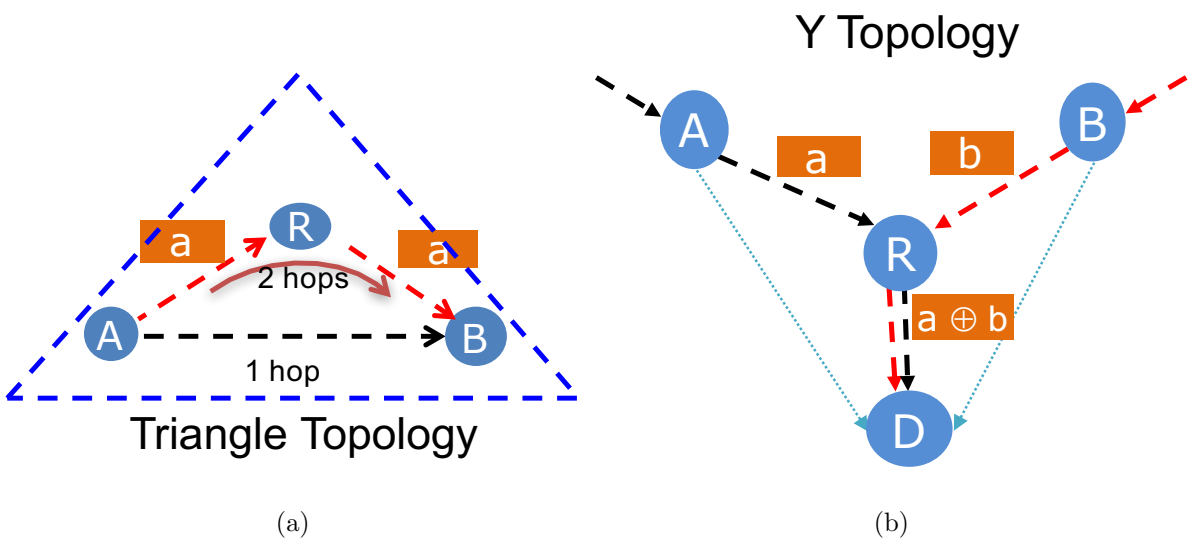


Figure 2.4: Types of topologies (a) Triangle topology (relay channel) (b) Y-topology (multiple access relay channel, MARC)

Butterfly (X) topology is depicted in Figure 2.5(a). In both Y and X topologies, the two data flows from node 'A' and node 'B' pass through an intermediate node 'R' to destinations 'D_1' and 'D_2' in Figure 2.5(a) and to a common destination 'D' in Figure 2.4(b). Node 'R' can get the chance to operate network coding function over the two incoming flows a and b and transmit the coded packets $a \oplus b$ to the destinations. The destination nodes have high chance to overhear the transmissions from the nearest senders without additional transmission cost and copies of those messages are useful for the recovery of

original messages. By this method, the intermediate node accomplishes the forwarding of received messages in a fewer number of transmissions and saves energy and bandwidth consumption.

In mesh topology Figure 2.5(b), there are direct connection links between every station. Wireless devices can utilise the advantage of short-distance communication links by creating a cooperative data sharing group among them via the built-in Bluetooth or Wi-Fi interfaces. It can be formulated as a network coding and transmission scheme. For example, some nearby wireless end devices such as mobile phones can benefit a fast downloading service from a common base station by cooperating to exchange their packets in mesh topological connections.

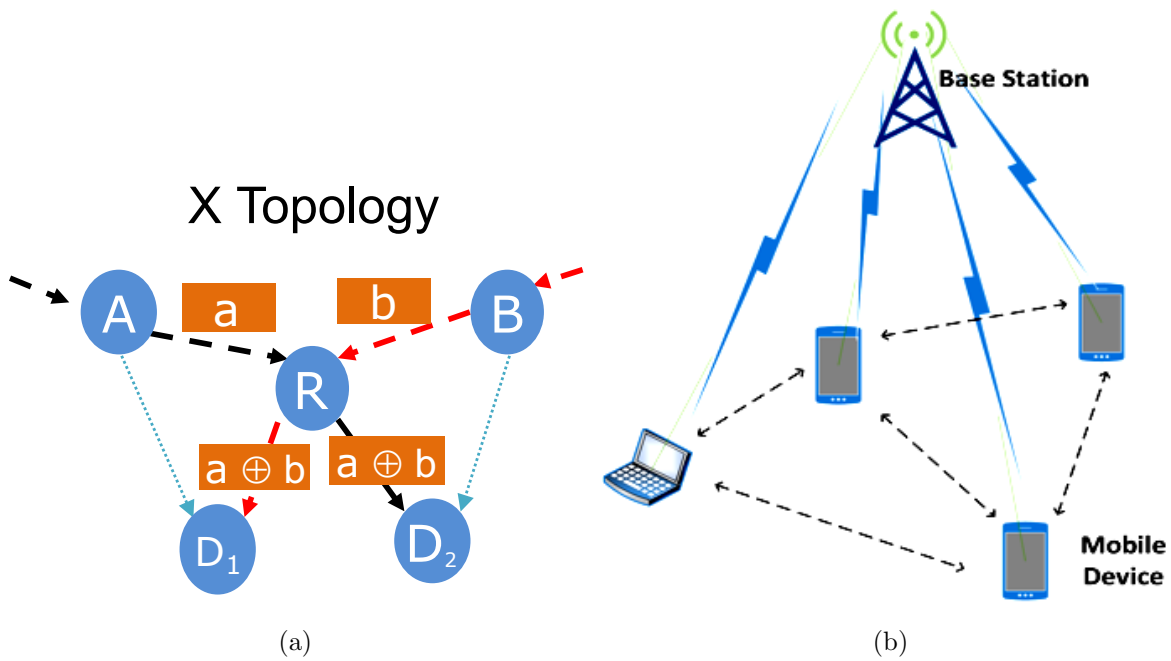


Figure 2.5: Types of topologies (a) Butterfly topology (b) Partially mesh topology

2.3 Multihop Wireless Networks (MWN)

In mulitihop wireless networks, nodes communicate with each other using wireless channels and do not have common infrastructure or centralised control. Any two nodes can communicate directly if their packets can be correctly decoded under the desired signal-to-interference-plus-noise-ratio [41]. In order to communicate with nodes beyond their

range of transmission, a wireless node has to depend on other intermediate nodes for relaying its messages to the desired destinations. Such an architecture requires that nodes in the network play the role of a source, a destination, or a router to relay the messages [42]. There are many branches of multihop wireless communication such as Mobile Ad hoc Networks (MANET), Wireless Sensor Networks (WSN), Wireless Mesh Networks, and Vehicular Ad hoc Networks with their own specific characteristics.

There are several benefits of MWNs such as extending the coverage of a network, improving the connectivity and capability of high data transmission rate when nodes are close to each other. MWNs also have important challenges. Determining a routing function for sending packets to the intended destination is one of the main challenges because of the characteristics of unreliable wireless links, packet loss and topology change. The routing protocol finds the path between the source and destination depending on the available links and intermediate nodes in the network. Long distance links can reach the destination in a few hops, but at some low speed.

On the other hand, the short distance links can support the transmission in high rate, but more hops are needed to reach to the destination. As a consequence, the chance of a node to be involved in relaying the other's data packets becomes high [43]. Therefore, the level of congestion at an intermediate node may become an important issue and lead to problems such as high energy consumption and buffer overflow of the node. However, this can be utilised as an opportunity for network coding to carry more information in one coded packet and possibly to reduce the required number of transmissions. The energy of some nodes is limited and the overhead for each single packet transmission is also important to be considered when multiple relays are involved for each active data flows.

2.3.1 Medium Access Control (MAC)

Among the services of link layer such as framing, reliable delivery, error detection and correction, link access is one of the services to move a datagram from one node to an adjacent node over a single communication link. A medium access control (MAC) protocol specifies the rules by which a frame is transmitted onto the link. For point-to-point links that have a single sender and a single receiver at each end of the link, the MAC protocol is simple. The sender can send a frame whenever the link is idle. However, when multiple

nodes share a single broadcast link, the MAC protocol serves to coordinate to avoid the collision. Typically, when there is a collision, none of the receiving nodes can recover the transmitted frames and all the frames involved in the collision are defined as loss. Therefore, the broadcast channel is wasted during the collision interval. In PLNC, this typical perspective is reversed. PLNC applies the advantage of signal level addition in the air and not all the collisions are assumed as loss. For the typical approach, MAC is necessary to coordinate the transmissions of the active nodes in order to ensure that the broadcast channel is not wasted when multiple nodes are active. It is responsible for resolving conflicts among different nodes for channel access. Since the MAC layer has a direct bearing on how reliably and efficiently data can be transmitted between two nodes along the routing path in the network, it affects the Quality of Service (QoS) of the network [44].

To achieve this, different kinds of techniques such as channel partitioning protocols and random access protocols were developed. Time-division multiplexing (TDM) and frequency-division multiplexing (FDM) are two techniques that can be used to partition a broadcast channel's bandwidth among all nodes sharing that channel. TDM divides time into time frames and further divides each time frame into N time slots. FDM divides the R bps channel into different frequencies (each with a bandwidth of R/N) and assigns each frequency to one of the N nodes. While TDM and FDM assign time slots and frequencies, respectively, to the nodes, code division multiple access (CDMA) assigns a different code to each node. Each node then uses its unique code to encode the data bits it sends. If the codes are chosen carefully, different nodes can transmit simultaneously and their receivers correctly receive a sender's encoded data bits [45].

2.3.2 Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA)

The CSMA/CA protocol is one of the two modes of the IEEE 802.11 MAC protocol, which uses the distributed coordination function (DCF) based on the RTS/CTS handshaking mechanism [44]. This protocol works in combination with a Binary Exponential Backoff (BEB) algorithm. Before sending data packets, a node first senses the medium for an idle channel in Distributed Inter Frame Spacing (DIFS) period. The node delays its

own transmission by a random backoff timer and waits for transmission to prevent a collision. A random backoff timer is chosen uniformly in the range of $[0, CW]$, where CW is referred to as the contention window. At the first transmission attempt, CW is set to the minimum contention window ($CW_{\min}$). When the timer times out and the medium is free, it broadcasts a Request-To-Send (RTS) control message, specifying a destination and data size. The receiver node responds with a Clear-To-Send (CTS) message. The control messages are transmitted after the medium has been free for Short Inter Frame Spacing (SIFS). If the sender node does not receive the CTS message, it may retransmit the RTS message. On receiving the CTS message, the sender sends the DATA and waits for an ACK from the receiver. Every node that hears the RTS/CTS exchange updates its Network Allocation Vector (NAV) value and must refrain from transmitting for that duration. When the NAV is set, the period for those nodes will remain busy until the end of ACK transmission. By using this method, all nodes including also hidden nodes can defer their transmission in an appropriate way to avoid collision. For each unsuccessful transmission, the value of CW is gradually doubled up to the maximum contention window value.

2.3.3 2-hop Path Selection Protocol (2PSP)

Lim and Yoshida in [46] developed a 2-hop Path Selection Protocol (2PSP) for a set of nodes, in which data can be sent faster using adaptive rate control capability of IEEE 802.11a/b/g MAC protocol via a relaying concept. Their main objective is to build upon the opportunistic rate adaptation to assist a sender, a relay and a receiver to reach a higher rate data transmission. They also proposed a relay mechanism and a new contention window called a Short Backoff Internal (SBI). A potential node that succeeds as a relay is allowed to send a Ready-To-Relay (RTR) message before transmitting payload data from the sender. The RTR message contains the information about a pair of higher transmission rates. These rates are determined by the relay depending on the signal quality of overheard Relay-Request-To-Send (RRTS) and Relay-Clear-To-Send (RCTS) messages from the sender and the receiver respectively. If the sender can decode the relay's message successfully, it sends data with a high data rate selected by the relay. If the sender cannot receive the RTR message, it can send the data according to the

standard DCF method. Then, the relay node forwards the received data to the receiver in another high data rate. Finally, the receiver sends the ACK message to the sender at the base rate. The DCF message handshaking procedure of 2PSP is depicted in Figure 2.6.

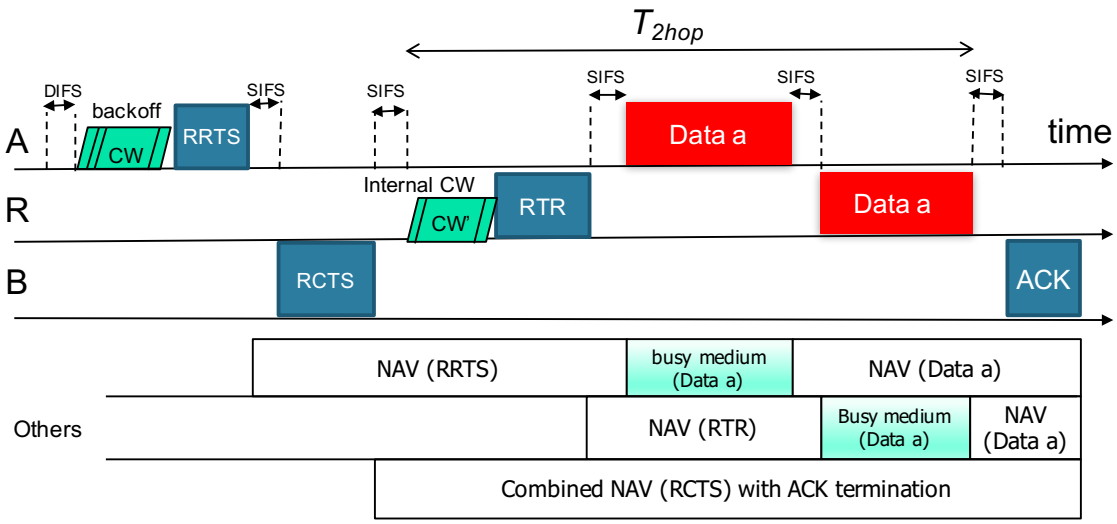


Figure 2.6: DCF message handshaking procedure of 2PSP protocol

2.4 Efficient Network Coding Based Data Transfer Framework (E-neco Framework)

A conceptual framework is developed for transferring data between wireless stations in wireless communication scenarios discussed in Section 1.2. In the near future, no technology will be able to provide alone the high demand of fast and reliable communication due to the very high increasing requirements of wireless and industrial communication. The network should be flexible in integration to work with other technologies. This framework is conceptualized based on the combination of the future potential network technologies such as D2D, M2M and MWN to support the new requirements. The multihop wireless communication is expected to become the core communication paradigm in future despite the fact that it is currently an extra option in IEEE 802.11. This framework is an important concept aiming for the future 5G because it depicts one of the very core parts of future wireless networks to provide the high-rate low-energy, fast and reliable services to

the billions of connected devices and machines. This framework is called E-neco framework, which refers to efficient network coding based data transfer framework. The general architecture of E-neco framework is depicted in Figure 2.7.

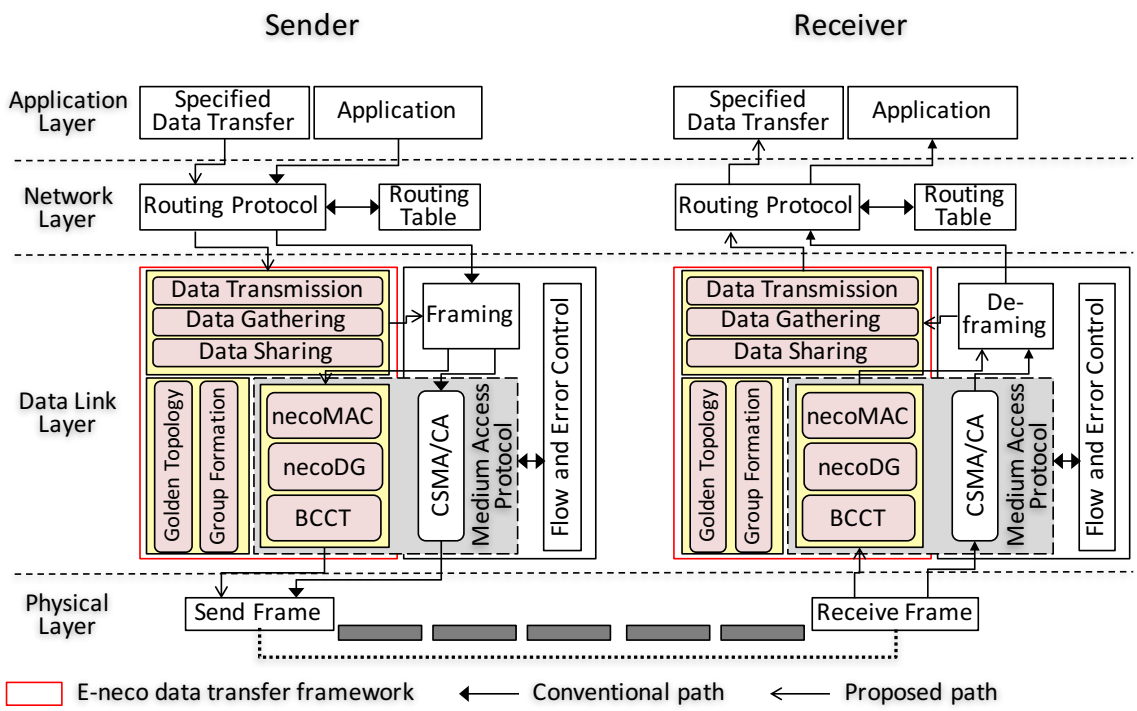


Figure 2.7: E-neco Data Transfer Framework

The framework utilises network topology and network coding techniques with MAC protocols. Therefore, the focus of this framework exists at data link layer. It consists of three main components:

1. data transfer mode
2. topology
3. medium access control protocols.

The functions of each component are briefly described as follows:

1. Data Transfer Mode: Three types of data transferring process can be accomplished by the framework. They are
    - Data transmission: Transmission of data from one node to another.

- Data gathering: Collection of data from many nodes, e.g., data gathering in a WSN.
- Data sharing: Distribution of possessed data to other members in a group for a common welfare.

The specific data transfer process is configured by the user at this moment. In future, a traffic pattern can be applied to activate the selection of the three modes.

2. Topology: Way of connecting links between nodes for network coding opportunity.

   - Golden topology: Types of topologies that creates network coding opportunity are called golden topology. They consist of Chain (also called linear) topology, triangle, diamond, Y-topology and cross (X) topology.
   - Group formation: Formation of logical connection among the participants in a group for data sharing purpose.

3. Medium Access Control Protocols: Rules by which a frame is transmitted onto the link. The proposed MAC protocols are designed to get maximum benefits from incorporation with network coding. They include necoMAC, necoDG and BCCT schemes for three data transfer modes.

   - necoMAC: NC-aware MAC scheme. This necoMAC scheme is intended for the data transmission in multirate multihop wireless networks. NC-aware 2-hop path selection protocol (NCA-2PSP) and NC-aware carrier sense multiple access (NCA-CSMA) protocols are proposed for golden chain and triangle topologies. The scheme also applies the multi-rate capability data transmission. High throughput, less energy consumption and low delay services are expected from this scheme.
   - necoDG: NC-based data gathering scheme for data collection from physical environment such as WSN. NC is applied at the CH or aggregator and at a relay inside a cluster of sensor nodes to assist reliable data transfer to the base station (BS). An incorporation of network coding-aware medium access control protocols and cluster-based WSN is made for the data transmission

from each sensor node to the aggregator or CH node. A modified 2PSP protocol is proposed to achieve more energy saving and longer network lifetime.

- BCCT: balanced cooperative coding and transmission scheme. A network coding-based cooperative data sharing scheme in mesh network topology. This scheme is proposed to satisfy the requirements of mobile data users downloading from a congested base station. Local group formation and network coded sharing is the main mechanism in this scheme. Participants are controlled by themselves using the reception information to maintain fairness and network lifetime.
- CSMA/CA: Carrier Sense Multiple Access with Collision Avoidance. The CSMA/CA protocol is one of the two modes of the IEEE 802.11 MAC protocol, which uses the distributed coordination function (DCF) based on the RTS/CTS handshaking mechanism [44]. This protocol works in combination with a Binary Exponential Backoff (BEB) algorithm. Before sending data packets, a node first senses the medium for an idle channel in Distributed Inter Frame Spacing (DIFS) period. The node delays its own transmission by a random backoff timer and waits for transmission to avoid a collision. The proposed scheme also works compatibly with CSMA/CA protocol.

These three components are relying on each other to perform a specific data transfer. Although they are depicted as separate components for easy understanding, in practice, their functions cannot be separated from each other. Depending on the desired data transferring process, the different topology and medium access control protocols are used. For example, BCCT scheme is selected for the balanced cooperative data sharing with the group formation mechanism of BCCT in mesh topology. The desired topology is created on the fly by the protocol. Therefore, topology and group formation can also be defined as part of the protocol. These two components have very close relation to each other. Topology and group formation are described as separate components of MAC protocols in the framework. This is to highlight the importance of topology for network coding benefit and thereby, for improving the performance of the system as a whole.

The core connection among these networks is the concept of multihop wireless communication. In every network scenarios above, network coding can be added as a function

of transmitter or receiver or forwarder at any point along the route depending on the information such as topology, data flows, route, energy etc. Although the purpose can be different from one target network to another depending on the application, the overall aim of the framework is to achieve the high performance from the operation of multihop communication. Whatever application scenarios come, the core operation of multihop communication and network coding functions can be applied. In other words, this framework is conceptualized for the future requirement of integration of network technologies.

2.5 Summary

This chapter presents some introduction to network coding and its background on different types of the network coding techniques and schemes. It also covers an explanation of multihop wireless communications and 2PSP protocol. The proposed framework of efficient network coding based data transfer is also described in this chapter.

Chapter 3

Network Coding-Aware MAC (necoMAC) Scheme

3.1 Introduction

In this chapter, a network coding aware Medium Access Control (necoMAC) scheme is proposed, and it incorporates many protocols such as Network Coding-aware 2-hop Path Selection Protocol (NCA-2PSP), 2-hop Path Selection Protocol (2PSP) and Network Coding-aware Carrier Sense Multiple Access (NCA-CSMA) in order to provide data transmission at higher rates with a fewer number of transmissions for multirate wireless networks. Two golden topologies called the golden chain and the golden triangle are mainly utilised for the proposed protocols. The energy consumption, overhead ratio, throughput and fairness for each protocol are calculated in the simulation. For the performance evaluation of the proposed protocols, a simulation program is developed and the performance of these protocols are analysed in many scenarios with varying numbers of nodes and flows. The simulation results reveal that the proposed scheme provides higher throughput and less energy consumption compared to the conventional Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA) protocol.

This chapter is organised as follows. Section 3.2 describes the problem statement and section 3.3 summarises the motivation derived from the existing works. Introduction to the two topologies called golden chain and golden triangle follows in section 3.4. Then, a description on how network coding can reduce the number of transmissions in a golden

chain topology and explanation of how it can be applied together with the 2PSP protocol is presented. The proposed necoMAC scheme and the new handshaking procedures of NCA-CSMA and NCA-2PSP protocols are described in Section 3.5. In this section, the brief explanation of the compared protocols, COPE and RD-DCF+NC are also described. Section 3.6 describes the analysis of protocols. In Section 3.7, the simulation parameters and scenarios are explained. The evaluation of the performance of proposed protocols is also performed with various metrics such as throughput, energy consumption, overhead ratio and fairness in different scenarios. Finally, Section 3.8 summarises the chapter.

3.2 Problem Statement

Energy consumption is an important issue that needs to be dealt with for a MWN such as Mobile Ad hoc Network (MANET) and Wireless Sensor Network (WSN) because the nodes in those networks rely on the limited battery capacity. Fennel and Nilsson [47] carry out a series of experiments to measure the energy consumption of an IEEE 802.11 wireless network interface operating in ad hoc networking environment. Their experiments show that fixed overhead costs are very high and improvements in data transmission rates have a fairly limited effect on overall per-packet energy consumption. This is due to the fact that the MAC protocol and broadcast traffic use lower transmission rates.

Current Medium Access Control (MAC) and physical layers for 802.11 based multihop networks impose high overhead for the transmission of small data packets [48]. The MAC functions are directly related to the reliable and efficient data transmission between two nodes along the routing path in the network [44]. As network coding [30] is a technique which can combine many receiving packets and forward the coded packet on the outgoing links, this problem can be mitigated by appropriate design of MAC protocol and network coding technique together. In [49], a network coding aware MAC protocol is proposed to combine batch transmissions and multiple reverse direction data exchanges. By this approach, any node in the network can transmit a burst of data packets in a single channel access invocation and high throughput is achieved.

Multihop relaying is a promising solution for increasing throughput and providing coverage for a large physical area with the cooperation of several intermediate nodes without the expensive physical infrastructure. The wireless medium is a shared resource, and

nodes that have packets to transmit access the medium in a distributed way, independently of their current traffic demand [48]. Researchers have been trying to improve the performance of MAC in many different ways, for example, to solve the inefficiencies of the base Distributed Coordination Function (DCF).

In [46], Lim and Yoshida create a 2-hop Path Selection Protocol (2PSP) which can reduce the delay and energy consumption by utilising the physical layer multirate capability with an introduction of a new Ready-To-Relay (RTR) message in the handshaking procedure of medium access.

In multirate wireless networks, data transmission can be achieved in different rates provided by the physical layer hardware. The conventional IEEE 802.11b and 802.11g/a specification provides up to 11 and 54 Mbps data rates, respectively [50]. Data transmission at higher rates can take place under favourable conditions as several environmental factors like channel condition and interference can have a dramatic impact on range and resulting coverage area. The receiver will not be able to decode the receiving message if the signal-to-interference-plus-noise ratio is lower than a certain threshold. If two nodes are at a short distance to each other, there is higher chance to transmit at higher rate because of the short link and strong signal quality as shown in Figure 3.1. IEEE 802.11b/g/a/n networks are currently the most popular wireless LAN products on the market.

With limited frequency resources, designing an effective MAC (Medium Access Control) protocol is a hot challenge [51]. As network coding has advantage of exploiting the broadcast nature of wireless transmission without extra cost, it has been introduced into the link layer to improve the performance of the network. Exclusive OR (XOR) network coding [52] is applied at the relay node to combine the incoming packets and to relay the

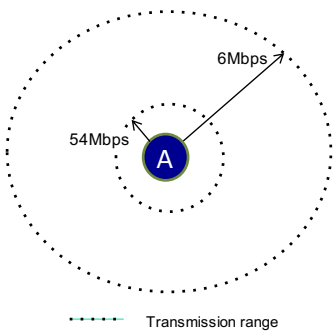


Figure 3.1: IEEE 802.11a ranges

coded packets in some higher rates.

3.3 Related Work and Motivation

The application of network coding into layer 2 has been done in [24]. In this work, network coding is integrated into the 802.11 infrastructure by adding a sublayer of network coding functionality between layer 2 and 3 of the seven layers of OSI reference model [53]. It combines layer 3 datagrams at access points (APs) and layer 2 broadcasting is performed. Layer 2.5 network coding works best in high rate symmetric traffic flows between 2 clients connected to the same 802.11 access point. Fairness becomes a problem in accessing the wireless medium and the services of AP due to the inefficiency of DCF in carrier sense multiple access with collision avoidance (CSMA/CA) protocol when the number of clients increase. Huang et al. [54] propose a combination of network coding with CSMA/CA protocol to enhance the fairness of wireless medium access among stations for a single-relay multi-user wireless network. They develop optimisation of the minimum contention window size according to the number of stations to create appropriate transmission opportunity.

Much research work on cooperative MAC with network coding is done in the literature. Research in this discipline involves the selection and coding functions relating to a relay node and exploitation of neighbour nodes to take advantage of more coding opportunities. BEND [36] is an example of MAC layer solution to practical network coding in multihop wireless networks. The advantage of BEND is its exploitation of the redundant copies of a packet in the neighbourhood coding repository and it allows the neighbour nodes to become a candidate coder and forwarder. Any node can code and forward a packet even when the node is not the intended MAC receiver that the packet is destined for. To further reduce the number of transmissions, a relay is allowed to retransmit the data for the source node together with its own data simultaneously in [37]. A novel Network Coding-Aware Cooperative Medium Access Control (NCAC-MAC) protocol is proposed based on IEEE 802.11 CSMA policy by further incorporating two collision free relay selection strategies to improve the relay selection process. PlayNCool [55] introduces the concept of a helper node in the neighbourhood of a sender and relay node. It applies RLNC technique for the network coded transmissions of packets. The helper node waits

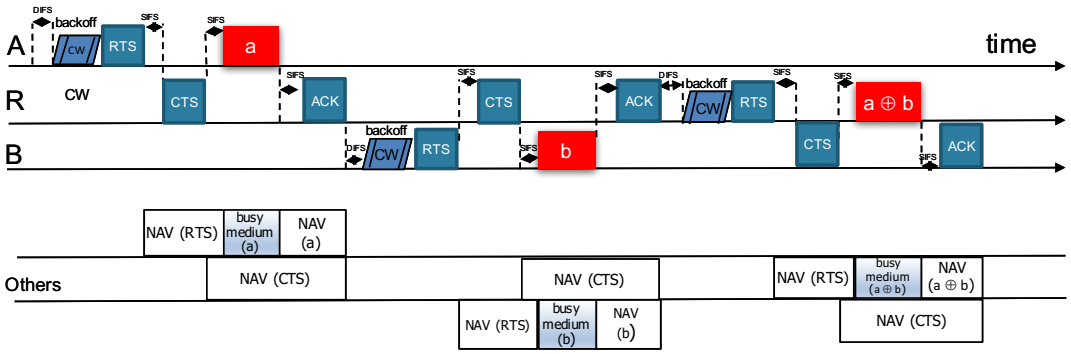
before transmitting the recoded packet while the medium accessing function is running. This approach can operate upon the path selected by the upper layer routing protocol and uses local link quality information for the helper selection. The great advantage is that PlayNCool protocol can operate in a variety of current routing protocols for the short-term deployment of network coding in real systems [20]. Some related works in the discipline of routing and network coding opportunity are TCAR [56] and ExCODE [57] for the local topology detection to decide the encoding of two packets with XOR coding technique.

The idea is to combine the property of network coding technique which can reduce the number of transmissions and the 2PSP protocol, which can exploit the multirate transmission capability of the IEEE 802.11 physical devices. The 2PSP protocol is designed for a single data flow between the sender and the receiver. The function of this protocol is incorporated with the XOR network coding technique over a golden triangle. The higher data transmission rate can be achieved by the 2PSP protocol, but this is only possible for the payload data. The overhead for control messages is constant for both small and large payload size because they are transmitted at lowest rate. With network coding, the entire one transmission of both overhead control messages and the data can be reduced by the coded transmission. Therefore, if a protocol that can exploit the benefits of these two techniques can be designed, it will be possible to save more energy, latency and thereby, achieve greater throughput improvement. However, the 2PSP protocol is designed to choose two-hop path instead of the direct one-hop transmission for a data flow; and XOR network coding is a technique that can be applied when two data flows meet at a node. Some questions are raised for motivating to do this research. They are "How to design a network coding based 2PSP to incorporate these two techniques?" and "What will be the performance of the created protocol?".

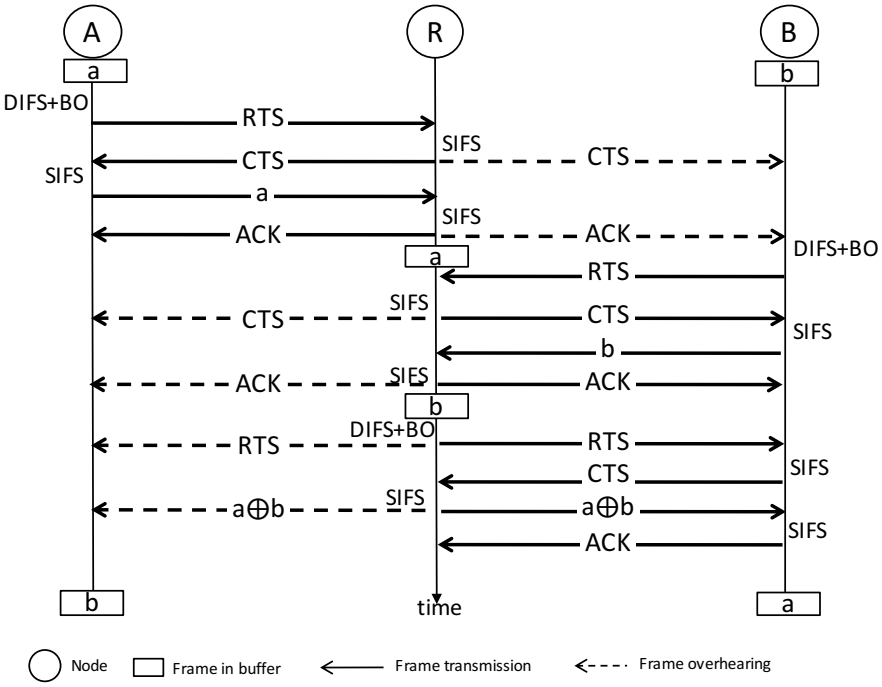
Reference COPE (DCF+NC) and RD-DCF with Network Coding

COPE combines the IEEE 802.11 DCF with the XOR network coding technique. It adds a layer of Network coding between the network layer and MAC layer. The relay node transmits the XOR-coded packets like in NCA-CSMA and reduces one transmission time slot. The MAC operation of COPE is according to the DCF of IEEE 802.11. The

relay node needs to compete for the channel access with the other nodes for transmitting the coded packet. The performance of NC is depending on the DCF, and network coding opportunity in COPE is limited by the DCF. COPE reduces the overhead control messages in its transmission cycle. It saves one transmission for the ACK message after receiving the coded packet by a source node. This ACK event will be added to the header of next coded packet to acknowledge the packet reception.



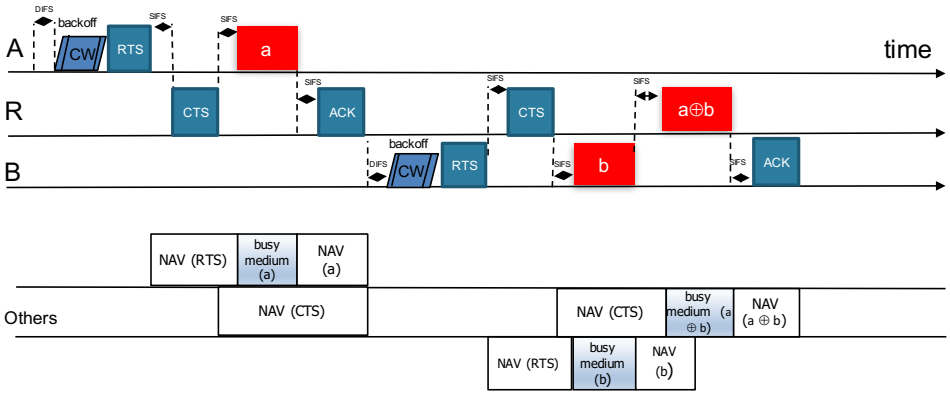
(a) Message handshaking with NAV



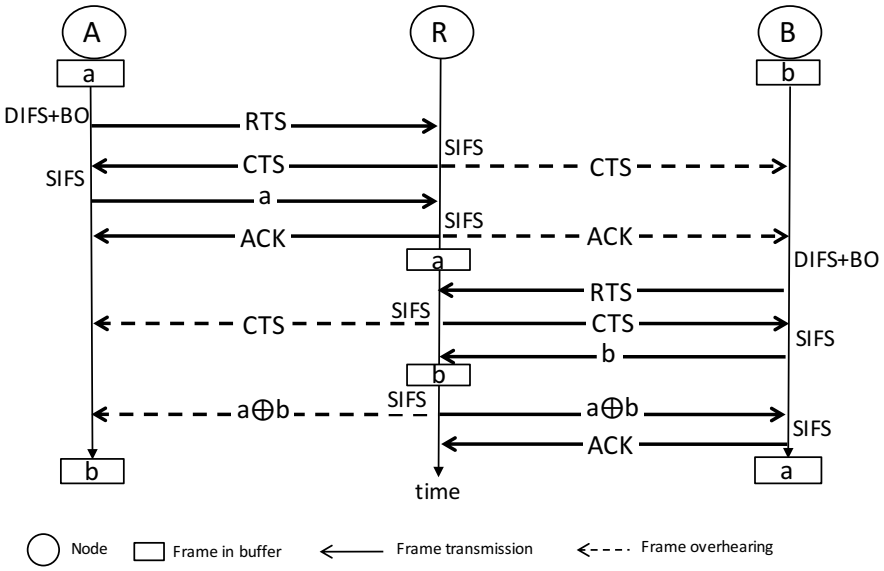
(b) Message handshaking timeline

Figure 3.2: DCF operation for data exchange in a golden chain for COPE

To solve the limitation of DCF on the network coding opportunity, [58] introduces a coding-aware MAC scheme with network coding. This scheme enables the reverse



(a) Message handshaking with NAV



(b) Message handshaking timeline

Figure 3.3: DCF operation for data exchange in a golden chain for RD-DCF with NC

direction (RD) communication between the relay node and any other station. Upon successful reception of a data packet the relay station can transmit a coded packet whose destination is the source of the received packet. The relay station can reduce the channel contention and the coded packet can be sent right after a packet is received. The value of the duration field in the transmitted data packet is extended to cover the channel access for the duration of transmission in reverse direction. Therefore, the throughput, delay and energy saving can be improved. Only when the channel condition is poor, this RD transmission will lead to the retransmission of both the forward and reverse data because the reverse transmission is used as an implicit ACK and the packet loss probability is higher for a data packet than for an ACK packet. The DCF MAC operation with the

RTS/CTS handshaking procedure for COPE and RD-DCF with Network Coding are presented for a golden chain topology in Figure 3.2 and Figure 3.3, respectively.

3.4 Golden Chain and Golden Triangle

In the necoMAC scheme, the design of the protocol together with the network coding operation are all dependent on the formation of golden topologies. The golden topology management block handles the golden chains and golden triangles inside the network. A golden chain is defined as a chain of three successive nodes with two data flows from opposite directions along the path to the destinations. Figures 3.6(a) and 3.6(b) show a golden chain. The data flow 1 and 2 form a golden chain at nodes A, R, and B, where A and B are outside the transmission range of each other.

For a golden triangle, A and B are within the 1-hop transmission range of each other. They may have a low-rate link between them. The main property of a golden triangle is the transmission in some higher rates is possible with the help of relay node. The difference from a golden chain is that a golden triangle works with one data flow or with two data flows in opposite directions. Figures 3.6(c) and 3.6(d) show the example of a golden triangle.

3.5 Proposed necoMAC scheme

Designing a MAC scheme that supports benefits of network coding means to provide more medium access for the coded packet transmissions than the other transmissions. It can be achieved by using the control messages and allocating the medium access for the encoder node (the relay) if network coding opportunity is detected. The reason is that due to the fairness mechanism applied in DCF of IEEE 802.11 the relay node cannot transmit the encoded packets at every network coding opportunity. The relay (the encoder) is assumed just as one of the transmitting nodes and it needs to compete for the medium access. Therefore, the destination nodes cannot recover the required packets as early as possible. In other words, network coded transmissions are limited by the MAC protocol. Therefore, MAC protocol needs to be designed to provide the benefits of network coding. In the proposed protocols, special control messages, Ready-To-Broadcast (RTB) and Ready-To-

Relay-Coded (RTRC), are introduced to achieve the purpose. The novelty of necoMAC exists at the combination of golden topologies, network coding and multirate transmission capability.

The proposed necoMAC scheme comprises data transfer mode, topology management and MAC protocols for data transmission. The scheme decides the most beneficial MAC protocol among NCA-2PSP, 2PSP, NCA-CSMA and CSMA/CA depending on the network topology and network coding opportunity of received packets. Figure 3.4 and Figure 3.5 show the proposed necoMAC scheme and its operation respectively. If a node detects two incoming data flows from opposite directions, it tries to decide whether the source nodes are within the transmission range of each other or not. If they can reach each other with one-hop transmission, the NCA-2PSP protocol tries to find a helper relay node and transmission will be performed with NCA-2PSP protocol. The transmission in NCA-CSMA will be selected when the source nodes are outside the transmission range of each other and they are located on a golden chain. These decisions can be described as a decision mechanism although the operations are not separately performed. In both cases the relay node decides a network coding opportunity if it detects two opposite data flows passing through it. Otherwise, if the golden triangle can be created with the help of a relay between the source nodes, data transmission in 2PSP is chosen. Data transmission in normal DCF mode will be perform if the above conditions do not match (in other words, the existence of only a single data flow without golden triangle topology).

If only one incoming data flow exists at a node, the scheme checks whether a golden triangle can be created or not. If the golden triangle can be created with the help of a relay, data transmission will occur in 2PSP. It will access the medium and send the data with the normal CSMA/CA protocol if both golden triangle and chain do not exist.

3.5.1 Network Coding-aware CSMA (NCA-CSMA)

Figure 3.6(b) depicts the data transfer with NCA-CSMA protocol in a golden chain topology. In this scenario, node A and node B are outside the transmission range of each other and within the transmission range of relay node R. A and B cannot transmit their packets to R in the same time slot to avoid collision at R. By the conventional CSMA/CA protocol, total 4 transmissions are needed to complete a successful data exchange between A

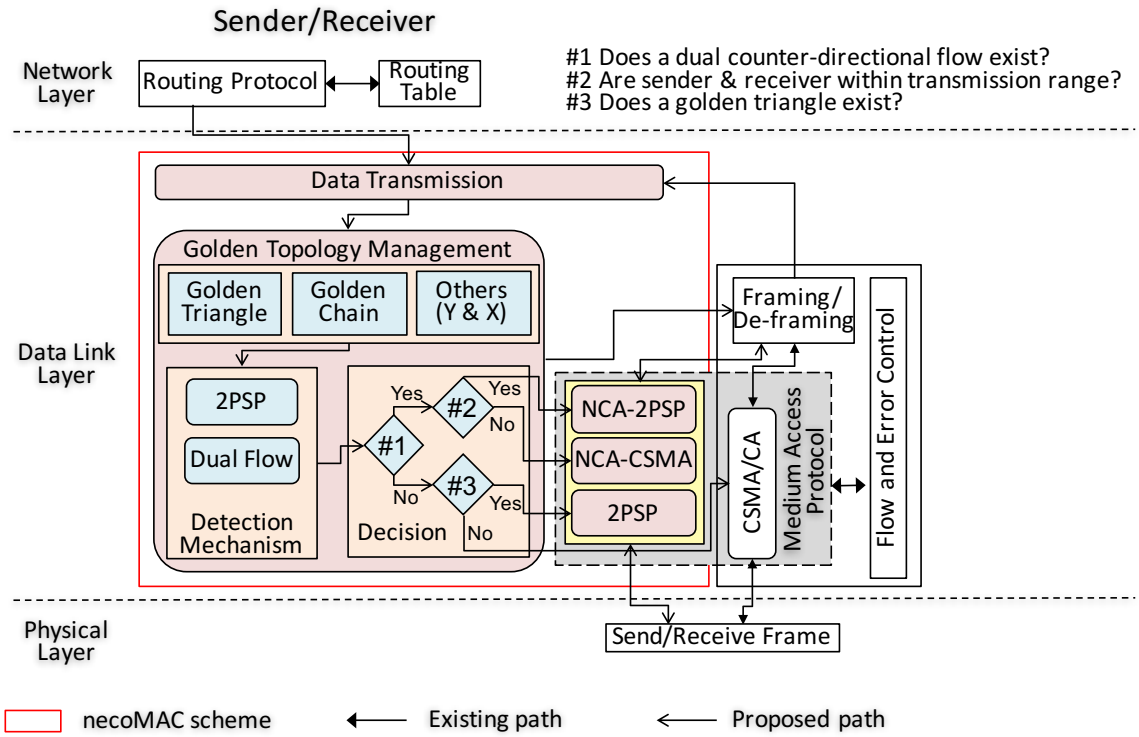


Figure 3.4: Block diagram of Network Coding-aware MAC (necoMAC) Scheme

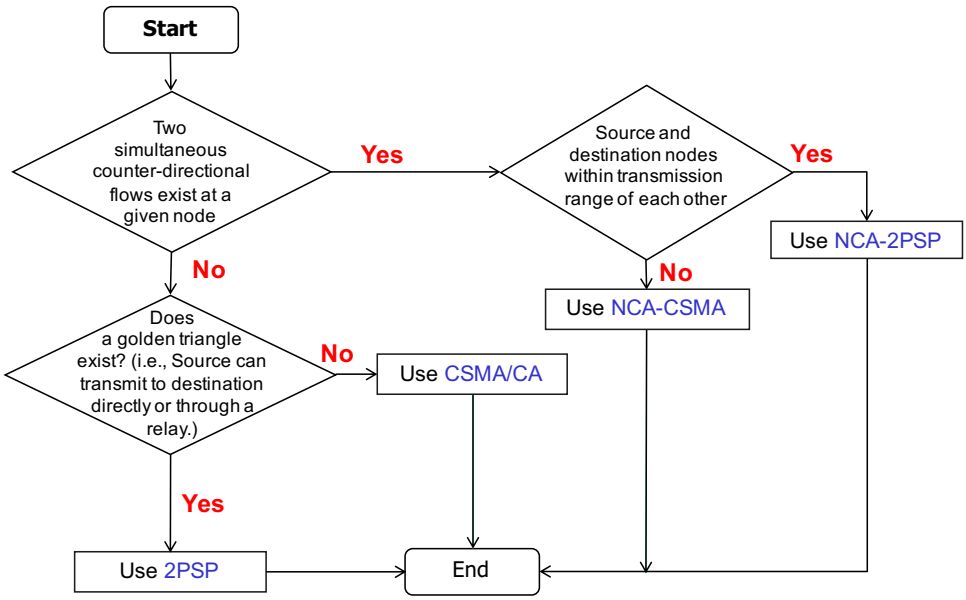
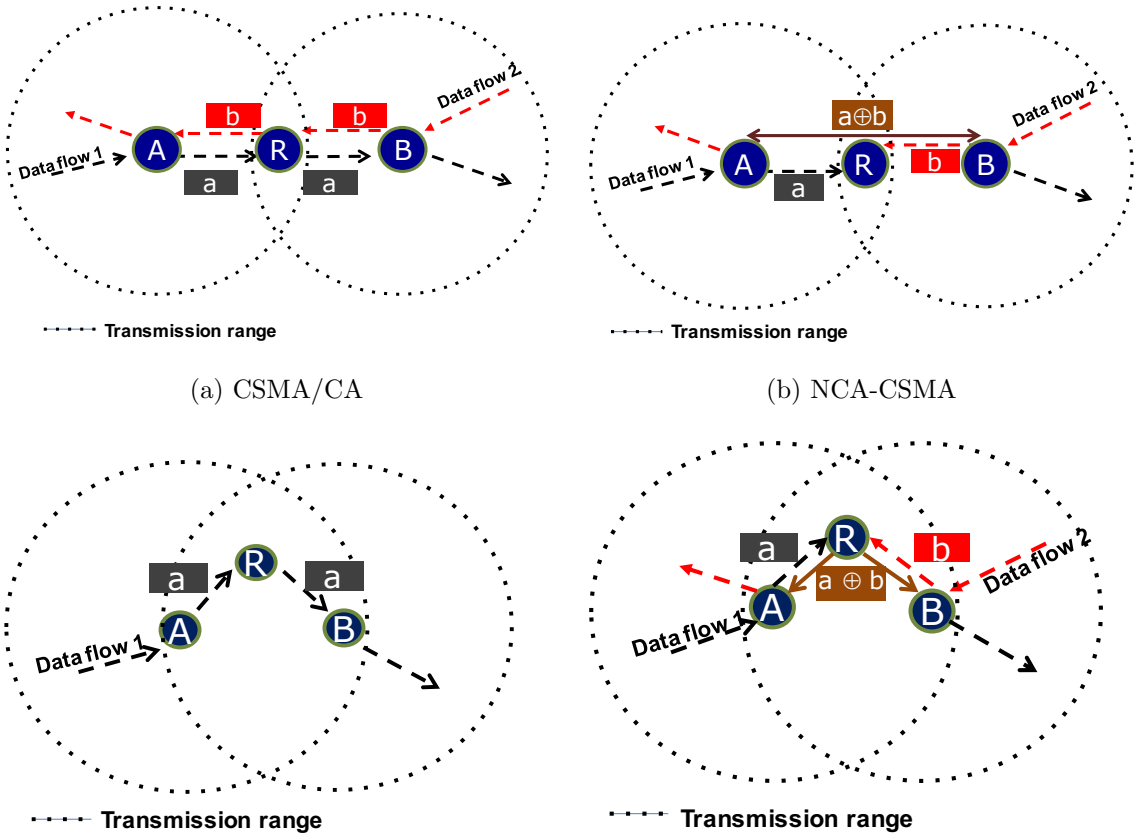


Figure 3.5: The Operation of necoMAC scheme

and B via the relay R, i.e., one transmission from each node A and node B and the two forwarding transmissions from the relay. For each transmission, a complete cycle of RTS, BO, CTS, DATA and ACK messages is needed in CSMA/CA. The message handshaking



(a) CSMA/CA

(b) NCA-CSMA

(c) 2PSP

(d) NCA-2PSP

Figure 3.6: Message exchange in reference scenarios: golden chain and golden triangle topologies

procedure of CSMA/CA protocol is depicted in Figure 3.7.

In the proposed NCA-CSMA, XOR network coding is applied at relay node R. The relay node R broadcasts the network coded packet $a \oplus b$ upon receiving the packets a and b from node A and B in previous time slots respectively. The relay node R notices that there is a network coding opportunity with packets a and b from two counter-directional data flows. Therefore, after successfully receiving b, the relay node immediately broadcasts a Ready-To-Broadcast (RTB) message to inform other nodes that it will broadcast a coded packet in next time slot. RTB can also be considered as an ACK to B and it is assumed that there is no packet loss due to channel errors and buffer overflow. RTB is designed like RTS frame format with two address fields.

Both A and B receive the XOR-coded packet in a time slot and can recover the packets b and a by performing XOR operation of their own packet with the receiving coded packet

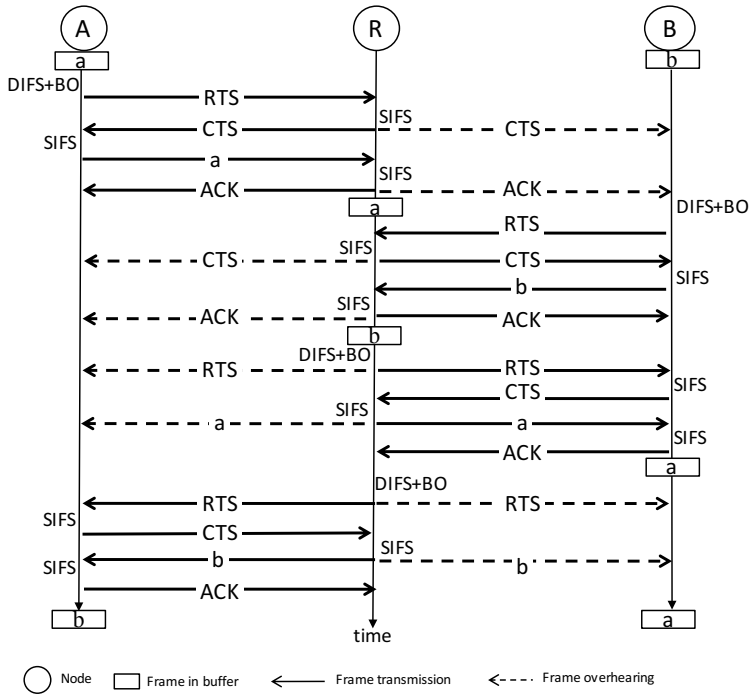
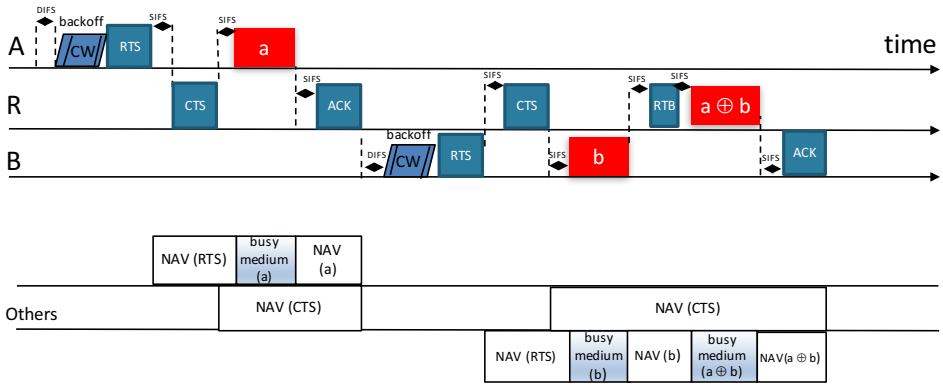


Figure 3.7: The CSMA/CA protocol message handshaking procedure for a data exchange between two nodes via a relay in a golden chain

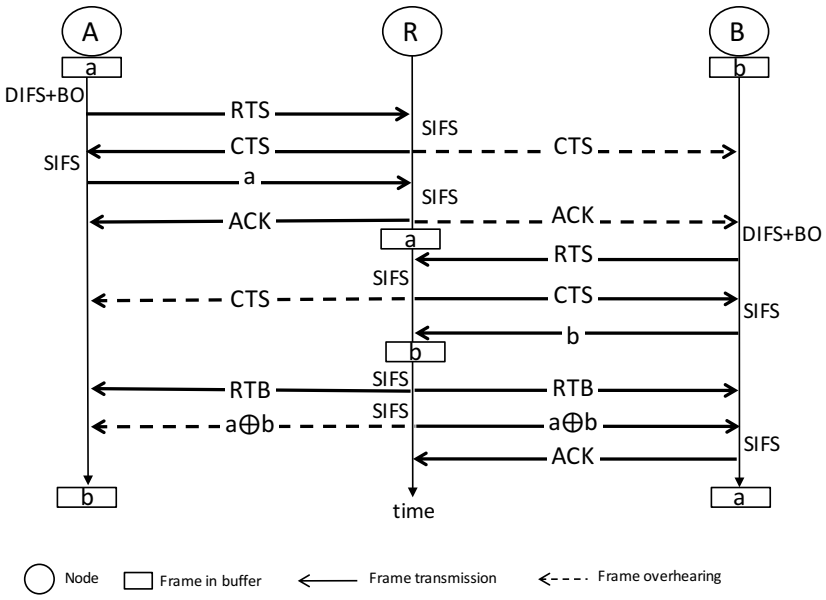
again. Therefore, data exchange between A and B is completed within three transmissions and energy for one transmission is saved. The procedure for messaging control and data packets is described in Figure 3.8. The difference between COPE and NCA-CSMA is that a separate ACK for B and a pair of RTS-CTS messages for $a \oplus b$ is used in COPE whereas NCA-CSMA uses RTB to replace these three messages. RD-DCF+NC [58] does not use RTB message for the transmission of coded packet. It directly allocates the channel access with reverse-direction transmission mechanism after receiving b.

The purpose of RTB is to reduce loss when coded packets are transmitted at high rates. It is intended to better perform than RD-DCF with NC in dense locations. RD-DCF with NC replaces the ACK for B with coded packet. Therefore, the transmission should normally be in low rate in order for the coded packet to be received successfully on both sides. In NCA-CSMA, RTB is used as an acknowledgement for B's data. Therefore, there is no limitation to transmit the coded packet only in low rate for successful acknowledgement. As the acknowledgement is satisfied by RTB's reception at B, the coded packet can also be transmitted at high rates if the condition is satisfied. Transmitting the coded packet at high rates is more beneficial as the size of coded packet is much bigger

than the ACK.



(a)



(b)

Figure 3.8: The NCA-CSMA protocol message handshaking procedure with NAV for a data exchange between two nodes via a relay in a golden chain

3.5.2 Network Coding-aware 2PSP Protocol (NCA-2PSP)

It is important to have right decision for network coded transmission and to have access the medium for the transmission of network coded packets after the decision, in order to exploit the full benefits of network coding-based transmission scheme. Therefore, network coding cannot be simply combined with 2PSP, a proper design of the NCA-2PSP protocol

is necessary to achieve the purpose. In this protocol, a control message called Ready-To-Relay-Coded (RTRC) is introduced. This function of this message is to invite transmission from the receiver for the network coding opportunity. RTRC control message carries information about the high transmission rates and network coding opportunity to the receiver and allows to transmit a packet after the transmission from the sender. If the receiver also has data to send, the medium is allocated for its transmission. Then, after performing the encoding operation, the relay node broadcasts the coded packet. Both the sender and the receiver can recover the desired packets after decoding operation. Finally, the receiver sends ACK to the sender and the operation is completed. If the transmission from the receiver is not detected, the relay will forward the packet received from the sender to the receiver as in 2PSP.

The NCA-2PSP protocol works when two nodes on a golden chain possess a helper relay between them and both of them have some data packets to be exchanged. For example, node A and node B in Figure 3.9 have some packets to transmit to each other and are waiting for some random backoff period. When the timer of node A times out earlier than B, it transmits a Relay-RTS (RRTS) message. If the receiver B receives the RRTS message correctly, it freezes its backoff counter and replies with a Relay-CTS (RCTS) message to the sender. At the same time, receiver B notices that it also has data packet whose destination is the source of the received RRTS message. Therefore, node B adds an information about its request to send in IL field in the RCTS message together with direct rate field and length field, which includes the duration for transmission both data a and b.

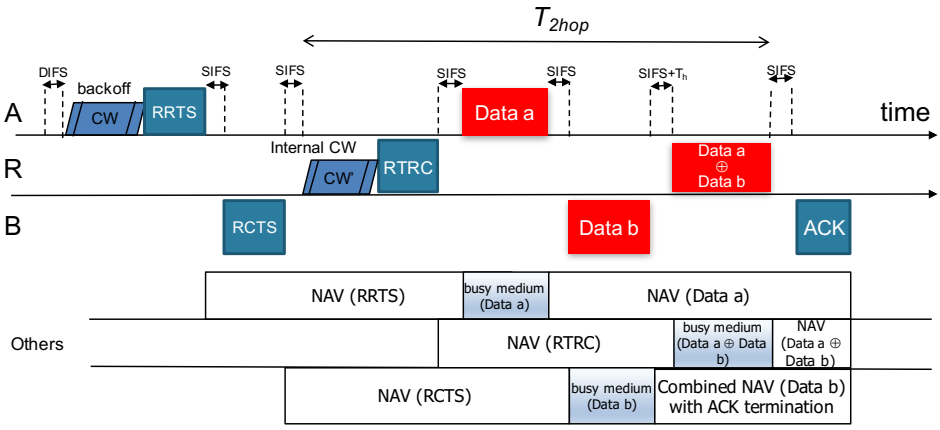
A relay node that hears these control messages determines a suitable pair of higher data rates based on the signal strength of the receiving RRTS and RCTS messages. The two chosen data rates must be higher than the 1-hop direct rate. The relay can compare these values because the information of direct rate is carried in the RCTS message. When the higher transmission rates have been decided, the relay node waits for a random Short Backoff Internal (SBI) period. The SBI timer is chosen uniformly in the range of $[0, CW^{'}]$, where the upper bound of $CW^{'}$ is equivalent to the minimum contention window CW_{min}. After the SBI times out, the relay broadcasts a Ready-to-Relay-and-Code (RTRC) message. This message contains the information about the two selected rates that the sender

and receiver should use when they send their data packets. Moreover, the length field of RTRC message is updated with the additional time required for the transmission of the coded packet $a \oplus b$ based on the information available in the RCTS message. Other nodes that overhear the RTRC shall update their NAV until the end of coded packet transmission.

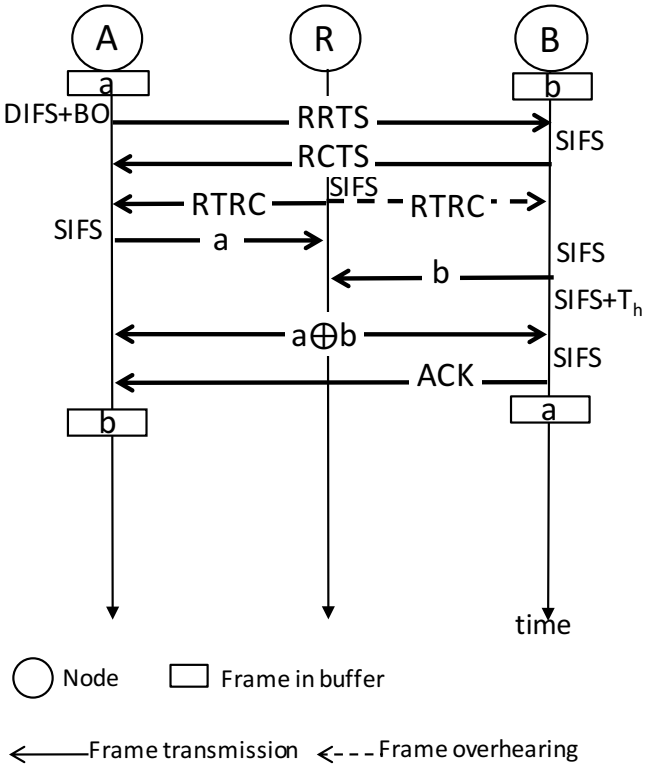
The detailed calculations for rate selection and relay mechanism are same to that of 2PSP. The NCA-2PSP is designed to backward compatible with 2PSP. If there is no data to send from node B when it receives RRTS, node B will simple reply with a RCTS message without any information about its desire to send a data packet like in 2PSP. If the sender cannot hear any response after SBI interval times out, the sender will transmit the data packet according to the standard DCF procedure. If the sender can correctly decode the RTRC message, it will transmit its data packet with the new data rate defined in the RTRC. The message handshaking procedure and control frame format of NCA-2PSP are depicted in Figure 3.9 and Figure 3.10 respectively.

A holding time of 10 ms is set for XOR coding operation after b is received at R. This value is obtained based on the analysis. After the relay node R receives the data packet of receiver B, it creates a XOR-coded packet of a and b, and then, broadcasts the coded packet. As both A and B are within one-hop transmission range of R, both of them can receive the coded packet and can recover a and b by doing the XOR function of their own packet and the receiving coded packet. After the data packets are successfully decoded, the ACK messages are sent. Neighbour nodes that hear the ACK shall terminate their NAV and are free for medium access. The handshaking procedure finishes when the sender receives the ACK message from the receiver node.

In NCA-2PSP, R can directly access the channel after data messages a and b are received while in COPE, R needs to compete the channel by sending RTS again. As compare to the RD+DCF with network coding protocol, NCA-2PSP can provide the coded packet transmission in multiple high rates with more reliability. This will be the main novelty as it can improve the performance to apply the advantage of network coding completely. As a whole necoMAC scheme, the novelty will be the decision of network coding opportunity and multirate transmission capability based on the information such as topology, routing information and the overheard signals from the neighbour nodes to



(a)



(b)

Figure 3.9: Message handshaking of NCA-2PSP Protocol in a golden triangle topology

achieve the maximum possible benefits.

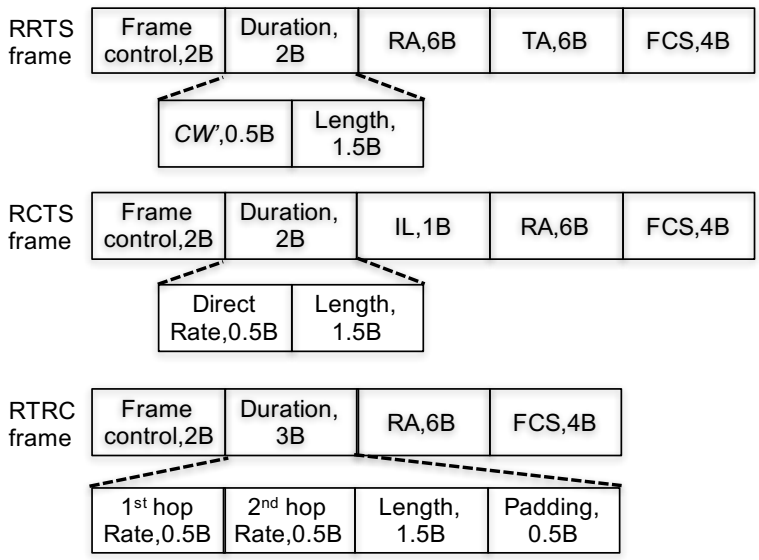


Figure 3.10: The Control Messages of NCA-2PSP Protocol

3.6 Analysis of Network Coding-aware MAC Protocols

To compute the upper bound of the throughput and energy consumption in idealistic conditions, the following assumptions are made. There are no packet losses due to channel errors and buffer overflow. The probability of collision is negligible. Encoding and decoding operations using XOR consume negligible time and energy. Fragmentation is not used. All nodes are equipped with IEEE 802.11a interfaces with omnidirectional antennas and each node can transmit, receive, or overhear data and control packets. The propagation delay is taken into account in the calculation.

The structure of a MAC data packet or MAC Protocol Data Unit (MPDU) consists of a MAC header, frame body or MAC Service Data Unit (MSDU), and Frame Check Sequence (FCS). When a MAC packet is to be transmitted, it is passed to the PHY layer and a PLCP Protocol Data Unit (PPDU) is formed. A PPDU consists of a PLCP preamble, a PLCP header and PLCP Service Data Unit (PSDU). The PLCP header except the SERVICE field is made up of SIGNAL field whose duration (T_{Signal}) is equal to the duration of a single OFDM symbol (T_{SYM}). The SERVICE field is included in the DATA with MAC header, MSDU and tail bits. The DATA part is changed into bits and divided by the Number of Data Bits per OFDM Symbol (N_{DPBS}). The resulting number

of symbols is multiplied by the duration of a symbol (T_{SYM}) to transform into time value. The N_{DPBS} values for each data rate are 24, 36, 48, 72, 96, 144, 192, and 216 respectively. The above parameters and their values are provided in Table 3.2. The delay to transmit each packet is calculated by the following equations. They are consulted with [59].

$$T_{RTS} = T_{Preamble} + T_{Signal} + T_{SYM} * \lceil \frac{L_{Service} + L_{Tail} + 8L_{RTS}}{N_{DPBS}} \rceil \tag{3.1}$$

$$T_{CTS} = T_{ACK} = T_{Preamble} + T_{Signal} + T_{SYM} * \lceil \frac{L_{Service} + L_{Tail} + 8L_{ACK}}{N_{DPBS}} \rceil \tag{3.2}$$

$$T_{DATA} = T_{Preamble} + T_{Signal} + T_{SYM} * \lceil \frac{L_{Service} + L_{Tail} + 8(L_{MAC} + MSDU + L_{FCS})}{N_{DPBS}} \rceil \tag{3.3}$$

Latency is the sum of transmission delay and propagation delay, $M/R + D$ seconds, where transmission delay is the time to put M-bit message "on the channel. M-bit is divided by data rate, R (bits/sec). Propagation delay, D, is the time for bits to propagate across the channel and it is calculated by dividing the length of packet in bits by speed of signals, c, which is the speed of light and equals to $3 * 10^8$ m/s. The length of a packet is 1500 bytes, and therefore, it is constant for every protocol. Delay for one transmission between two nodes is calculated by (3.4) and that for a pair of data exchange in a chain is as in (3.5) for CSMA/CA. The equation (3.5) and those for COPE and RD-DCF+NC are taken from [58]. Equations (3.6), (3.7) and (3.8) are derived for NCA-CSMA, 2PSP and NCA-2PSP respectively. T_{C-DATA} means the duration for transmission of XOR-coded packet and N is the number of data packets. The values shown in Table 3.1 are calculated as a sample with $N = 2$ data packets for each model shown in Figure 3.6. For the analysis purpose, the highest data rates are used to compute the maximum and minimum values for each metric.

$$Delay = T_{DIFS} + 3T_{SIFS} + T_{BO} + T_{RTS} + T_{CTS} + T_{DATA} + T_{ACK} \tag{3.4}$$

$$\begin{aligned} Delay_{CSMA/CA} = & \frac{2N}{N}(T_{DIFS} + 3T_{SIFS} + T_{BO} + T_{RTS} + T_{CTS} \\ & + T_{DATA} + T_{ACK}) \end{aligned} \tag{3.5}$$

$$Delay_{NCA-CSMA} = \frac{1}{N}[N(T_{DIFS} + 3T_{SIFS} + T_{BO} + T_{RTS} + T_{CTS} + T_{DATA} + T_{ACK}) + \frac{N}{2}(T_{RTB} + T_{C-DATA})] \tag{3.6}$$

$$Delay_{2PSP} = \frac{N}{N}(T_{DIFS} + 5T_{SIFS} + T_{BO} + T_{RRTS} + T_{RCTS} + T_{RTR} + 2T_{DATA} + T_{ACK}) \tag{3.7}$$

$$Delay_{NCA-2PSP} = \frac{1}{N}[N(T_{DATA}) + \frac{N}{2}(T_{DIFS} + 7T_{SIFS} + T_{BO} + T_{RRTS} + T_{RCTS} + T_{RTRC} + T_{C-DATA} + 2T_{ACK}) + \frac{N}{2}(T_{RTB} + T_{C-DATA})] \tag{3.8}$$

Table 3.1: Theoretical values of proposed protocols and the legacy systems for MSDU length 1500 bytes and data rate 54Mbps in golden chain and triangle topologies

Protocol	**Topology**	**Latency (ms)**	**Throughput (Mbps)**	**Overhead Ratio**	**Energy (μJ)**
CSMA/CA	Chain	0.921	13.029	0.2073	92.1
NCA-CSMA	Chain	0.601	19.983	0.06	60.1
2PSP	Triangle	0.755	15.905	0.0794	75.5
NCA-2PSP	Chain + Triangle	0.514	23.335	0.0466	51.4
COPE	Chain	0.695	17.272	0.0625	69.5
RD-DCF+NC	Cross	0.593	20.236	0.0502	59.3

Throughput is defined as the number of bits contained in a payload data, MSDU divided by the time required to transmit the data packet that includes the payload data. Throughput is calculated by (3.9) for all protocols, where Delay represents the delay for each specific protocol.

$$Throughput = \frac{MSDU\ (bits)}{\text{Delay}} \tag{3.9}$$

Energy consumption is the total energy consumed for the transmission of control messages and data packets. The energy consumption for one transmission between two nodes is the multiplication of transmission power and time taken from (3.4). The energy consumption of a pair of data exchange between two nodes in a golden chain is calculated by multiplying the transmission power and the delay for CSMA/CA, NCA-CSMA, 2PSP and NCA-2PSP protocols from (3.5), (3.6), (3.7) and (3.8) respectively. The assumption is that all nodes transmit at the same transmit power and only busy time is considered.

The overhead ratio, OR, is defined as the ratio of the total number of control messages (in bytes) and the total number of control messages plus the total number of payload bytes to transmit a data packet. The overhead ratio for the exchange of two data packets in a chain is formulated as follows for each protocol:

$$OR_{CSMA/CA} = 4 * \frac{RTS + CTS + ACK + L_{MAC}}{RTS + CTS + ACK + DATA + L_{MAC}} \tag{3.10}$$

$$OR_{NCA-CSMA} = \frac{2(RTS + CTS + 2ACK) + RTB + 3L_{MAC} + L_{XOR}}{2(RTS + CTS + DATA + 2ACK) + RTB + CDATA + 3L_{MAC} + L_{XOR}} \tag{3.11}$$

$$OR_{2PSP} = 2 * \frac{RRTS + RCTS + RTR + ACK + 2L_{MAC}}{RRTS + RCTS + RTR + ACK + 2DATA + 2L_{MAC}} \tag{3.12}$$

$$OR_{NCA-2PSP} = \frac{RRTS + RCTS + RTRC + 2ACK + 3L_{MAC} + L_{XOR}}{RRTS + RCTS + RTRC + 2ACK + 2DATA + CDATA + 3L_{MAC} + L_{XOR}} \tag{3.13}$$

$$OR_{COPE} = \frac{3RTS + 3CTS + 4ACK + 3L_{MAC} + L_{XOR}}{3RTS + 3CTS + 4ACK + 2DATA + CDATA + 3L_{MAC} + L_{XOR}} \tag{3.14}$$

$$OR_{RD-DCF+NC} = \frac{2RTS + 2CTS + 2ACK + 2L_{MAC} + L_{XOR}}{2RTS + 2CTS + 2ACK + 2DATA + CDATA + 3L_{MAC} + L_{XOR}} \tag{3.15}$$

The comparison of proposed protocols and the legacy systems is described in Table 3.1.

3.7 Numerical Simulation

In this section, the performance of the proposed necoMAC scheme is investigated compared to the conventional CSMA/CA scheme, RD-DCF+NC and COPE protocols. Some assumptions are made for the simulation. Firstly, all nodes are equipped with IEEE 802.11a hardware and work in multihop communication manner. The second assumption is that there is no packet loss due to channel errors or buffer overflow, and that coding and decoding using XOR operation consume negligible time and energy. Fragmentation is not used in the simulation. All nodes have one pair of (data rate, control rate) among (6,6), (9,6), (12,12), (18,12), (24,24), (36,24), (48,24), and (54,24). The performance of proposed protocols is evaluated with various metrics such as throughput, energy consumption, overhead ratio and latency. The results of the necoMAC scheme incorporate NCA-2PSP, NCA-CSMA, 2PSP and CSMA/CA protocols. The scheme decides a suitable protocol to transfer the data packets depending on the opportunity of network coding and golden topologies. If any of these conditions does not exist, the scheme will perform its operation in basic CSMA/CA mode. Likewise, the results of COPE and RD-DCF+NC also include CSMA/CA as a basic operation mode in the implementation for the conditions where no golden chain exists.

3.7.1 Simulation Scenarios and Parameters

Simulation scenarios are categorised into two. In the first scenario, the investigation of the impact of increasing number of data flows on protocol performance is performed with the fixed number of 40 nodes. The number of data flows is varied from 10 to 22 flows in steps of 4 flows. In the second scenario, the focus is on the influence of the number of nodes on the performance of the protocols. The number of data flows is fixed at 14 flows, and the number of nodes is varied from 35 nodes to 50 nodes in increment of 5. The simulation program is created in MATLAB environment and all nodes and flows are uniformly distributed in a $1000 \times 1000\ m^2$ coverage area. Source and destination pairs are randomly selected with the setting of maximum and minimum hop count 4 and 2 respectively because the proposed models are for 2 hops. The routing paths between them are produced in the similar way to the AODV routing protocol. All data flows are equal in size (1500 bytes) and one unicast data flow for each path. The proposed

scheme utilises the three-node golden chains and golden triangles along the paths to the destination for each flow. In each run of simulation, all nodes, source-destination pairs and corresponding data flows are newly generated. The simulation is run over 500 times and simulation results are averaged.

3.7.1.1 Verification of the Simulator

The results of the simulator are verified by generating a small scenario and comparing with the theoretical reference values obtained from the analysis as done in [59, 60]. All nodes are assumed to be equipped with the IEEE 802.11a radio interfaces in the multihop wireless network communication. The IEEE 802.11a hardware specifications are used for the simulation as shown in Table 3.2. Firstly, a pair of random source and destination is generated on the network coordinates. And the theoretical maximum throughput (TMT) for transmission from source to destination is calculated with the varying MAC Service Data Unit (MSDU) sizes from 500 bytes to 4000 bytes. For this simulation, the following assumptions are made. Bit error rate (BER) is zero. There are no losses due to collisions. The MAC layer does not use fragmentation. The graphs for 802.11a OFDM is presented in Figure 3.11 for RTS/CTS. The curves are plotted for the data rates of 6Mbps, 12Mbps, 24Mbps and 54Mbps. The graphs produced by the own simulator show the same performance as in the reference paper. The number of data bits per OFDM symbol N_{DBPS} values are 24, 48, 96 and 216 for OFDM-6Mbps, OFDM-12Mbps, OFDM-24Mbps and OFDM-54Mbps, respectively. Throughput is calculated as the amount of information contained in MSDU divided by the time required to transmit the data packet that includes the MSDU from source to the destination. TMT and the total delay per MSDU are calculated by the following formulae.

$$TMT = \frac{\text{MSDU size}}{\text{Delay per MSDU}} \tag{3.16}$$

$$\text{Delay per MSDU} = T_{DIFS} + T_{BO} + T_{RTS} + T_{CTS} + T_{DATA} + T_{ACK} + 3 * SIFS \tag{3.17}$$

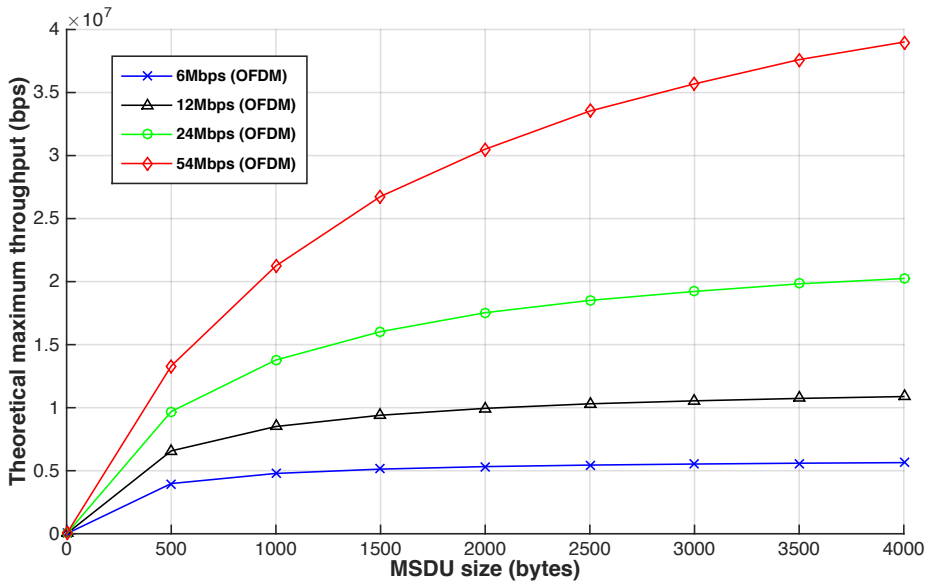


Figure 3.11: Theoretical maximum throughput curve for 802.11a OFDM RTS/CTS. The MATLAB simulator shows same performance results as in the reference.

3.7.1.2 Selection of Interference and Path Loss Model

The interference and path loss models are added to the simulator. When it comes to choosing an interference and loss model, an accurate interference model is very important because there could be many simultaneous transmissions in multihop wireless networks. The author in [61] points out that the model people usually used for the wireless network simulation is too simple and the results obtained are far from real conditions. In wireless networks, signal transmissions are in broadcast nature and suffer from interference of each other. In a wired network, links are connected between two nodes which are wired and the operation of each link has no interference with other links. However, the wireless networks have different limitations from the wireline networks. The wireless signals attenuate over distance due to path loss, shadowing and time-varying fading.

The transmission rates are calculated based on the SINR value obtained from all the interference of other nodes' transmissions. The additive interference model as in [61] is applied in the simulation. In this model, packet collisions are determined by the cumulative interference and noise instead of using the interference from a single node, one at a time. The success or failure of a single transmission is determined by the bit error rate (BER) and the signal-to-interference-plus-noise ratio (SINR) has to exceed an

appropriate threshold.

The SINR perceived by the receiver of link m, γ_m is calculated as follow.

$$\gamma_m = \frac{P_m \mathcal{A}(A_m, B_m)}{\sum\limits_{l \in L, l \neq m} P_l \mathcal{A}(A_l, B_m) + N_m f} \tag{3.18}$$

P_m is the received power. $\mathcal{A}(A_m, B_m)$ is channel attenuation from point A_m to location of the receiver B_m of link m. N_m and f denote the power spectral density of the thermal noise at the receiver of link m, and the frequency bandwidth of the channel, respectively. If SINR value is greater than the SINR threshold corresponding to an acceptable bit error rate (BER), the transmission is successful and the receiver receives the packets. Otherwise, there is a collision. SINR threshold 10 dB and Additive white Gaussian noise (AWGN) are used in the simulation. The power received at the receiver antenna is calculated by the Friis transmission formula. This is one of the fundamental equations in antenna theory. This formula relates the free space path loss, antenna gains and wavelength to the received and transmit powers. The formula can be described based on the λ value

$$P_r = P_t G_t G_r \left(\frac{\lambda^2}{(4\pi R)^2} \right) \tag{3.19}$$

or in frequency

$$P_r = P_t G_t G_r \left(\frac{c^2}{(4\pi R f)^2} \right) \tag{3.20}$$

, where P_t is the output power of transmitter antenna. G_t and G_r are the gain of the transmitting and receiving antenna, respectively. λ is wavelength in meter, R is the distance in meter between the transmitter and receiver. f is frequency and c is the speed of light (299,792,458 m/s).

3.7.2 Performance Metrics, Results and Discussion

3.7.2.1 Throughput

Figure 3.12(a) and Figure 3.13(d) respectively show the throughput per flow values of CSMA/CA, necoMAC, COPE and RD-DCF+NC protocols with the increase in the number of data flows and nodes. The throughput per flow increases with the increasing number of data flows for all protocols except CSMA/CA and it decreases with the increasing number of nodes for all protocols. This is due to the high opportunity of network coding over

Table 3.2: Parameters of PHY and MAC Layers for 802.11a

Parameter	**Value**
Transmission power	100 mW
RTS, RRTS size	20 bytes
CTS, ACK size	14 bytes
RCTS, RTR size	15 bytes
RTB size	21 bytes
MAC header	30 bytes
MAC header1	28 bytes
MAC header2	24 bytes
FCS size	4 bytes
XOR header	40 bytes
Slot time	9 μs
Preamble time	16 μs
Signal time	4 μs
SYM time	4 μs
DIFS	34 μs
SIFS	9 μs
L_{Tail}	6 bits
$L_{Service}$	16 bits
Average Backoff time (BO)	67.5 μs
Holding time T_h	10 ms

golden chain topology for many data flows in a fixed network with 40 nodes. The lines for necoMAC, COPE and RD-DCF+NC clearly show this point. For CSMA/CA, although data flows increase, it does not improve the throughput due to its DCF fairness. On the

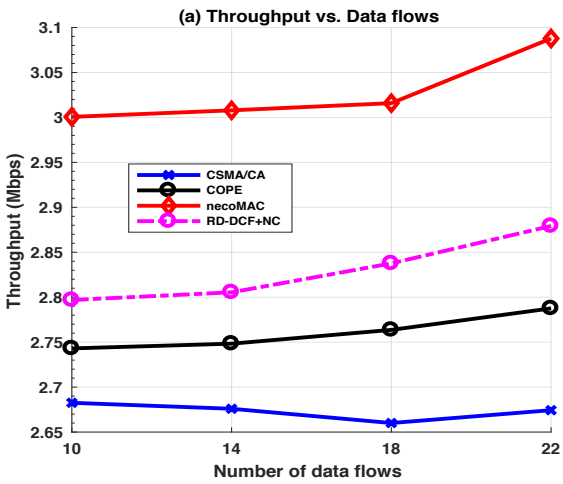
other hand, throughput decreases with the increasing number of nodes for all protocols. This is because only existing many nodes has no great effect without the active data flows. The number of flows is fixed at 14 for this experiment. For that value, throughput improves until 35 nodes and gradually decreases when the number of nodes increases. With some network coding opportunity and golden chain, other protocols have higher throughput than the CSMA/CA despite the similar trend. The similar or three-hop long. The necoMAC is the best among the four because it includes 2-hop path selection mechanism in the scheme which provides advantage of utilising relay-based transmissions in the network with many nodes. As the number of nodes increases in the network, the opportunity of finding a helper relay is higher. In both cases, necoMAC has 11% higher throughput per flow than the CSMA/CA protocol and 9% than COPE. There is a trade-off between the increasing number of flows and nodes. The NCA-2PSP protocol can find more relay nodes that help transmissions in high rate and network coding operation can further reduce the number of transmissions, which lead to the improvement in throughput when the number of flows and nodes increase. Only a few flows in a network with many nodes cannot achieve the full advantage.

3.7.2.2 Latency

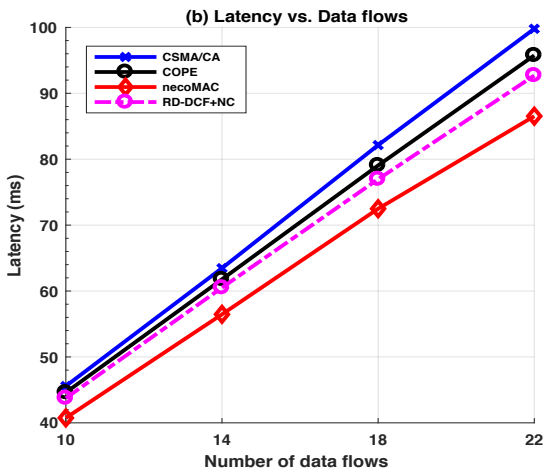
Figure 3.12(b) and Figure 3.13(e) depict the latency value versus the number of data flows and the number of nodes, respectively. The necoMAC scheme can gain an average of 11% delay reduction over the CSMA/CA protocol for each case. The value of latency becomes high when the number of data flows increase in the network with 40 nodes. For necoMAC, latency is 40 ms for 10 data flows and 86 ms for 22 flows. The necoMAC is 7% less latency than RD-DCF+NC and 8% less than COPE. For the network with increasing nodes, high latency means low throughput for constant payload size and flows size.

3.7.2.3 Energy Consumption

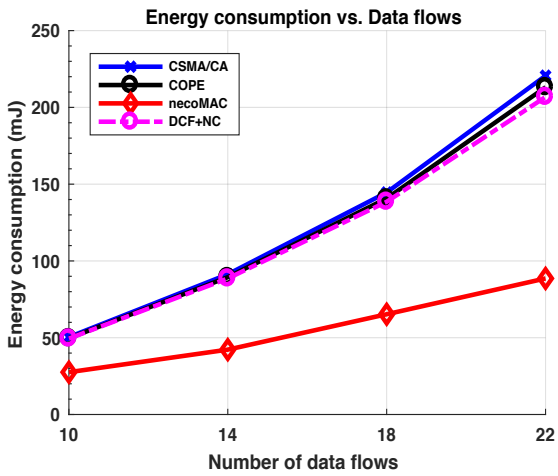
Figure 3.12(c) and Figure 3.13(f) show the energy consumption of protocols under evaluation. The total energy consumption increases as the number of data flows and nodes in the network increases according to the increase in latency for both cases. However, the necoMAC scheme consumes significantly less energy than the other protocols. The



(a) Throughput vs. Data flows



(b) Latency vs. Data flows

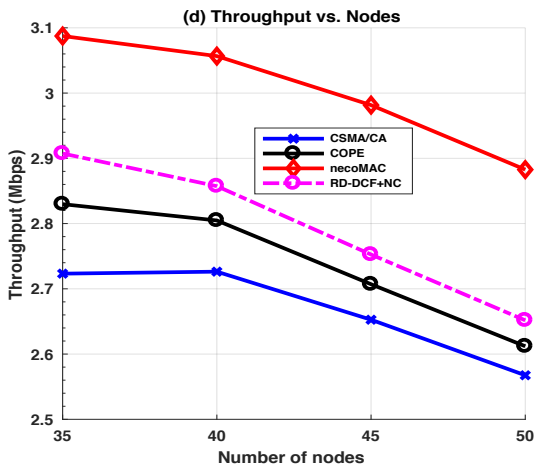


(c) Energy Consumption vs. Data flows

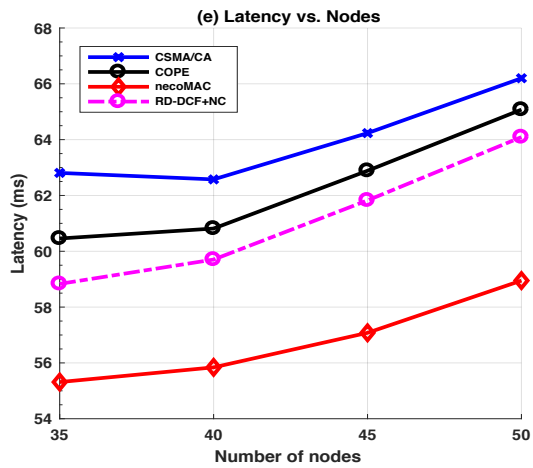
Figure 3.12: Throughput, latency and energy consumption of the necoMAC, CSMA/CA, COPE and RD-DCF+NC protocols as a function of increasing number of data flows in the golden chain and golden triangle topological networks

amount of energy saving is, on average, about 60% for the case of increasing flows and about 58% when the number of nodes increases. The reason for this decrease is, with the fixed number of 14 flows, increasing number of nodes create more opportunities to find a relay node between the sender and receiver. This enables the data transmission at the high rates and in a shorter duration of the transmission, which leads to less energy consumption. Other protocols do not include the 2-hop path selection mechanism and take longer time to transmit a packet to the receiver compared to the necoMAC scheme.

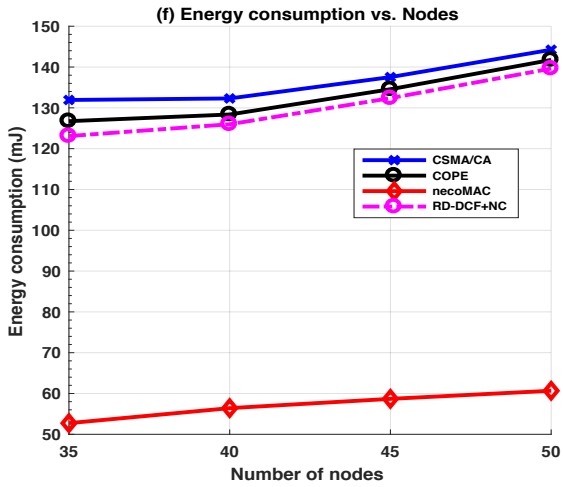
The necoMAC scheme chooses a relay based on the rate with direct transmission and another achievable higher rate with two-hop transmission like the work of [46].



(a) Throughput vs. Nodes



(b) Latency vs. Nodes



(c) Energy Consumption vs. Nodes

Figure 3.13: Throughput, latency and energy consumption of the necoMAC, CSMA/CA, COPE and RD-DCF+NC protocols as a function of increasing number of nodes in the golden chain and golden triangle topological networks

3.7.2.4 Overhead Ratio (OR)

The results of overhead ratio are for the exchange of a pair of data packets in a chain topology and are taken from Figure 3.1. The value of one means that the size of the

control messages is equal to the size of the control messages plus data payload. It should be desirable to transfer a certain amount of payload data with the smaller amount of control packets. The overhead ratio of the proposed NCA-2PSP protocol for payload size 1500 bytes is 0.0466, which is the smallest among the compared systems. Another proposed protocol, NCA-CSMA possesses an overhead ratio of 0.06, which is a little higher than the RD-DCF+NC with 0.0502 and a little smaller than COPE (0.0625). CSMA/CA protocol has the highest value of overhead ratio with 0.2073 and 2PSP is the second highest with 0.0794. These two protocols are the only ones that do not have network coding mechanism in their protocol designs. The observation from these values describes that the protocols which utilise the network coding opportunity can save time slots and reduce transmissions of control packet. This leads to the overall reduction in overhead for the completion of data exchange in a golden chain topology. The NCA-2PSP has 77.5% improvement than the CSMA/CA protocol and 7% than the state of the art.

3.8 Summary

This chapter proposes a network coding aware medium access control scheme called necoMAC, which incorporates NCA-CSMA, NCA-2PSP, 2PSP and CSMA/CA protocols to exploit the network coding techniques at the MAC layer. The network coding-aware CSMA is applied at a three-node golden chain and network coding-aware 2PSP works in a golden triangle with multirate transmission capability. The scheme achieves advantages for both cases of increasing the number of data flows and nodes in the network. Increasing the data flows creates more opportunity for golden chain and network coding while increasing the number of nodes provides more chance to search a relay node between source and destination nodes. The necoMAC scheme shows higher throughput, low latency, low energy consumption and low overhead radio than the existing systems. The proposed scheme provides 9% high throughput, more than 50% less energy consumption and 7% less overhead ratio than the state of the art systems.

Chapter 4

Network Coding Based Data Gathering (necoDG) Scheme

4.1 Introduction

This chapter presents the use of network coding techniques in the data gathering process of wireless sensor networks (WSNs), which possess the limited energy and computing resources to achieve the low latency and high reliable communication. The traditional way of reliable communication is based on the retransmission of packets and acknowledgements. It costs bandwidth and high latency for retransmissions of massive data from the sensor nodes. Energy consumption and the number of transmissions are used as the performance evaluation metrics.

The rest of this chapter is organised as follow. In Section 4.2, the problem statement is described. Section 4.3 presents the different kinds of data gathering methods in a WSN for energy efficiency and reliability. Section 4.4 describes the proposed scheme, which comprises the network architecture, energy-efficient 2-hop path selection protocol, and the construction of random linear network coding technique. The simulation and results of the proposed system are discussed in Section 4.5. Finally, Section 4.6 concludes the chapter.

4.2 Problem Statement

Wireless sensor networks (WSNs) can be defined as a network of spatially distributed autonomous sensors (often hundreds) that sense the physical or environmental properties such as temperature, pressure, humidity, lighting, sound intensity, vibration to monitor the desired physical process in the environment. The acquired information is wirelessly communicated to a base station (BS), which propagates the information to the remote devices for storage, analysis, and processing. For many scenarios, direct communication between two peers is limited due to the distance or the obstacles between them. Multihop communication is a desirable property for a WSN and participants in a WSN works in multihop fashion. Being a participant in a multihop network, a node is responsible for many tasks such as finding a next hop to send the data, routing decision and forwarding the data received from a neighbor node. All nodes in a WSN except the BS have their primary job, sensing the physical environment and transmitting the acquired information. These jobs consume the available energy of each sensor node. The massive deployment of sensors and actuators is a typical application for IoT. Connectivity and longer battery life are requirements for the collection of updated data from the physical environment. The ultra low latency and ultra high reliability are needed for the future application scenarios.

4.3 Related Work and Motivation

4.3.1 Data Gathering Methods in WSNs

For the durable battery life of individual sensor node or the longevity of the whole WSN, energy-efficient operations are needed. The related works existing in the literature for this purpose can be classified into three main groups. The first group focuses on the in-network data aggregation. Providing the summary of the data collected by computing the statistical metrics (e.g., average, sum, max) can reduce the amount of data to be transmitted from a large number of sensors to a distant sink [62, 63]. Another approach of this same group is the data compression based on the compressive sensing theory [64], in which the sink receives only a few encoded data of all the readings from sensors, being able to recover the original data, and reduces the global communication cost. These

two approaches are called in-network aggregation with size reduction and in-network aggregation without size reduction in [65] respectively.

The second classification of energy-saving operations for a WSN is the design of routing protocols to route the packets from different sources directing towards the same destination. They are different from the traditional ad hoc routing protocols in the metrics they use to select the paths. For example, in data-centric routing, sensor nodes should route packets to the next hop based on the packet content to promote in-network aggregation [65]. Shortest path routing is also a technique in which a path with the least hop counts is selected to reduce, for example, the delay. Routing protocols in WSNs differ from each other depending on the application and the network architecture. Furthermore, they can also be classified into multipath-based, query-based, and QoS-based protocols.

The third group of classification is the structure of nodes working together as a tree, or a cluster, or as a centralised approach. It may be considered as part of the routing protocol design because they are directly related to each other. Tree-based routing usually bases on the hierarchical organisation of the nodes in the network. A spanning tree is constructed rooting at the sink and aggregation is performed level by level from its leaves to the sink by answering the queries generated by the sink [66]. The drawback of this approach is due to the fact that wireless channels are not reliable. The loss at a given level of the tree means losing all data from that subtree. On the other side, a cluster-based approach has many advantages.

A number of clustering algorithms have been designed for the WSNs in the literature depending on the network architecture or the characteristics of the cluster head nodes [67, 68]. A cluster head may be elected by the sensors in a cluster or pre-assigned by the network designer. The cluster head can be a member of other sensors or a special sensor with high energy and computing resources. It can also be a member of a second-tier network or just a neighbour of the base station. Clustering can save the communication bandwidth as the member sensors of a cluster only communicate to the cluster head and avoid interactions with other cluster. Because the cluster head is already defined, sensors will only need to take care of connecting to their own cluster head and the overhead used for topology maintenance is reduced.

Although many data aggregation techniques and routing protocols are designed for

a number of performance metrics such as energy efficiency, reliability, quality of service, etc., there is an important issue relating to the energy efficiency of the entire process of data gathering in WSNs. It is the error and loss in wireless communication due to the dynamic channel conditions.

A larger percentage of energy is especially consumed for the transmission than for sensing and computation in sensor nodes. Therefore, it is important to provide reliable communication to improve network lifetime of wireless sensor networks. The traditional way to provide the reliability is to use the feedback messages to report the received or lost packets. However, these feedback messages consume bandwidth. The main objective of retransmission scheme is to improve the throughput but at the expense of energy consumption. Energy efficient error control techniques to prolong network lifetime in resource-limited network like WSNs and wireless communication remains a challenge [69]. Proper choice of error control code to provide the reliable communication and making the system energy efficient is another smart way.

Network coding can indirectly reduce the amount of bandwidth resources for the transmissions and improve the reliability for transmissions from the cluster head (CH) to the base station (BS). The transmissions from the CH to BS can be achieved by many available communication technologies such as Wi-Fi, ZigBee, 2G/3G/4G and it is also expected to support by coming 5G with the cost of very high power consumption. LPWA technologies can provide low power consumption and support large networks with low data rates.

In the proposed data gathering scheme, network coding is applied at the cluster head for the transmission of coded packets to the BS, which is the destination of many sensor nodes over a low-rate and long-link to improve the reliability of transmissions. The transmissions from the sensor nodes to CH is routed by an energy-efficient 2-hop path selection protocol to transmit over a high-rate link and reduce delay and energy consumption. This transmission can also achieve more connectivity among the sensor nodes.

4.4 Proposed Energy Efficient Data Gathering Scheme

The proposed scheme is for the data gathering process at a sink node, which is the destination of all the other sensor nodes in the network. Especially this research focuses on the data transmission process from the sensor nodes to the sink. The data aggregation

functions on the acquired information is not considered in a cluster head or at the sink. The transmitted data packets may be the raw data packets or the aggregated ones. The basic idea is to provide reliable communication for the data transmissions from sensor nodes to the sink, because the sink is very far from the sensor nodes. The proposed network coding-based data gathering scheme is depicted in a block diagram in Figure 4.1.

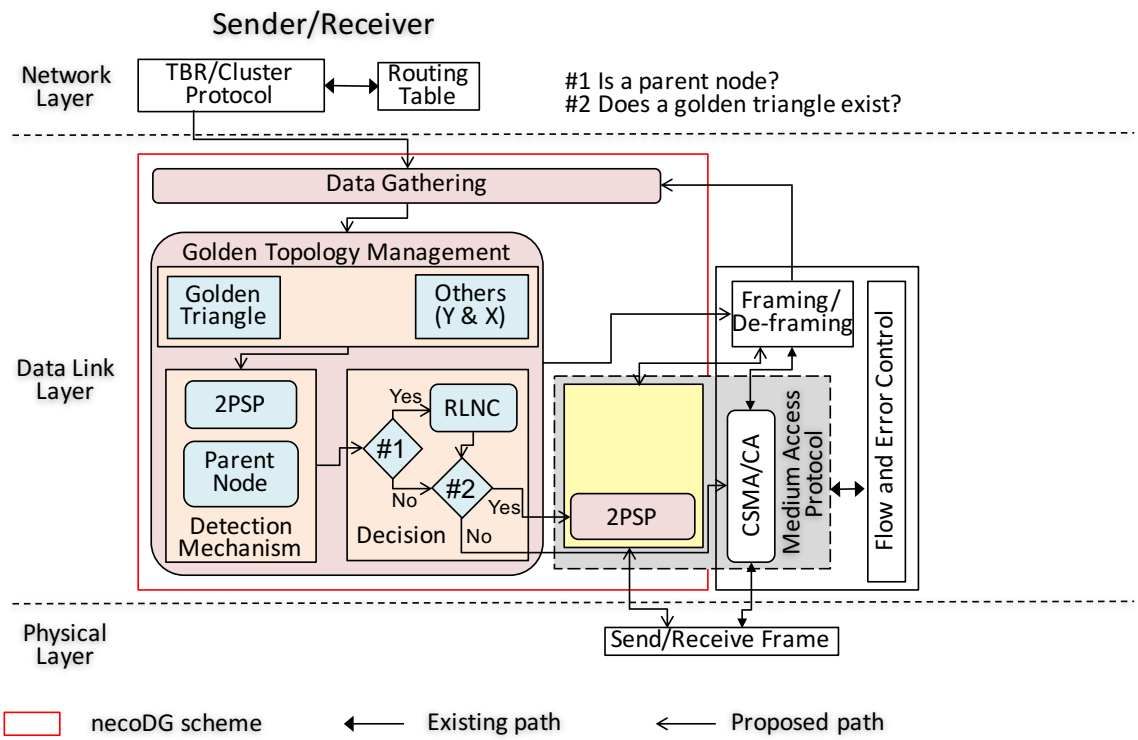


Figure 4.1: Block diagram of Network Coding-based Data Gathering (necoDG) Scheme

4.4.1 Network Architecture for the Proposed Scheme

Since the transmission from a sensor to a far BS is usually in low speed, the time taken to reach a long distance will be high. Therefore, clustering approach is selected in our work to make the data transfer more efficient and reliable. The problem in a WSN in multihop communication fashion is that the nodes around the sink have to relay (forward) more packets from other nodes inside the network. This may burden them and lead to their energy depletion early. The selection of CH is performed before the data transmission phase in a round. Therefore, new CH may be selected in every round depending on its remaining energy. The CH becomes a leader node to perform the random linear network

coding function over the incoming data packets from the member sensor nodes of its own cluster.

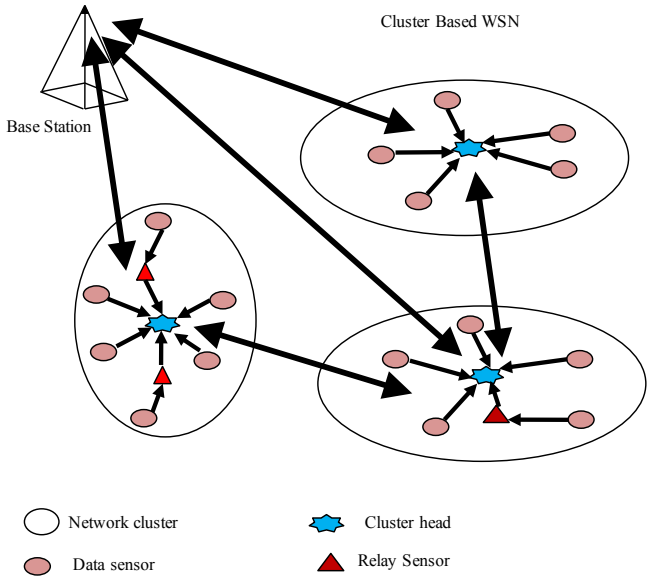


Figure 4.2: Data gathering in a cluster-based wireless sensor network (WSN)

The cluster-based architecture consists of a base station (sink), which is connected to an infrastructure network with high resources, and one or more cluster heads representing their own clusters of sensor nodes. The number of clusters depends on the clustering process of the clustering algorithm like LEACH [70] and the number of sensor nodes inside the network. LEACH (Low Energy Adaptive Clustering Hierarchy) is a protocol architecture for microsensor networks that combines the ideas of energy-efficient cluster-based routing and media access together with application-specific data aggregation to achieve good performance in terms of system lifetime and latency. The cluster heads can directly communicate to the sink or through a relay node in a multihop fashion. They are relay nodes that forward the sensor data from the cluster head to the sink.

In each round, a cluster setup phase is performed to select cluster heads and their related clusters. After that phase, data transmission from the sensor nodes is allowed. In LEACH, sensor nodes transmit their data to the CH in TDMA. In the proposed scheme, TDMA is replaced by the energy-efficient 2PSP in multihop transmission approach. The advantage of using multihop communication is the better connectivity among the sensor nodes. At the later round of the algorithm, there will only be small number of sensor nodes as their energy is depleted round by round. Connectivity and longer battery lifetime

are requirements for the collection of updated data from the physical environment.

4.4.2 Network Coded Data Transfer from Cluster Head to BS

The reliability of the transmissions from the CH to BS is important because they represents the data from the whole cluster of sensor nodes in a physical region. Therefore, CH takes the important role for the data forwarding between the sensor nodes and the BS. Another role of cluster heads is that they perform the network coding to merge the packets coming from different sources into the same packet so that a packet of the same size can carry much more physical information. These coded data packets are transmitted to the sink.

According to the technique of RLNC, more redundancy is included in the transmitted coded packets, which will improve the reliability of communication, and the overhead for feedback messages due to packet error and loss are reduced. The reason for choosing RLNC is that the algorithms for coding and decoding are well understood. The size of the encoded packet is the same as the size of an incoming packet, but an encoded packet generally carries information about several original packets. The combination of many incoming packets from different sensors is performed in a cluster head. The number of packets being coded together will be different from time to time depending on the packets received at a cluster head of the sensor nodes.

4.4.3 Modified 2-hop Path Selection Protocol (M2PSP)

An energy-efficient 2-hop path selection protocol is proposed based on 2PSP for local transmission within cluster to achieve high rate transmission and low energy consumption. The original 2PSP is capable of transmission in high rate by the link layer multi-rate capability. The reason of preferring the higher rate is to accomplish data transmission in low delay and improve the throughput. The focus in the proposed scheme is to provide reliability and to achieve the longer network life. Therefore, the relay selection mechanism selects a relay node with low energy consumption for forwarding the data message. To achieve this, the original 2PSP is modified in relay selection decision by introducing new rules to the MAC DCF to achieve data transmission through the relay in both high rate and low energy consumption. The energy consumption is estimated using the achievable

rates from the RTR mechanism.

As in 2PSP described in Section 2.3.3, the three new control messages called relay RTS (RRTS), relay CTS (RCTS) and Ready to Relay (RTR) are introduced. When a sender wants to send a data packet, it first waits for the DIFS+BO time period if the channel is busy before it can transmit the data. After the sender can access the idle channel, it transmits a RRTS message. If the receiver receives the RRTS message correctly, it replies a RCTS message to the sender. A relay node that hears these control messages waits for a random short back off internal (SBI) period. Within this period, the relay node decides whether to help the sender or not by estimating the energy consumption of direct transmission and the transmission with its help. The timing operation procedure with the modified 2PSP protocol is depicted in Figure 4.3.

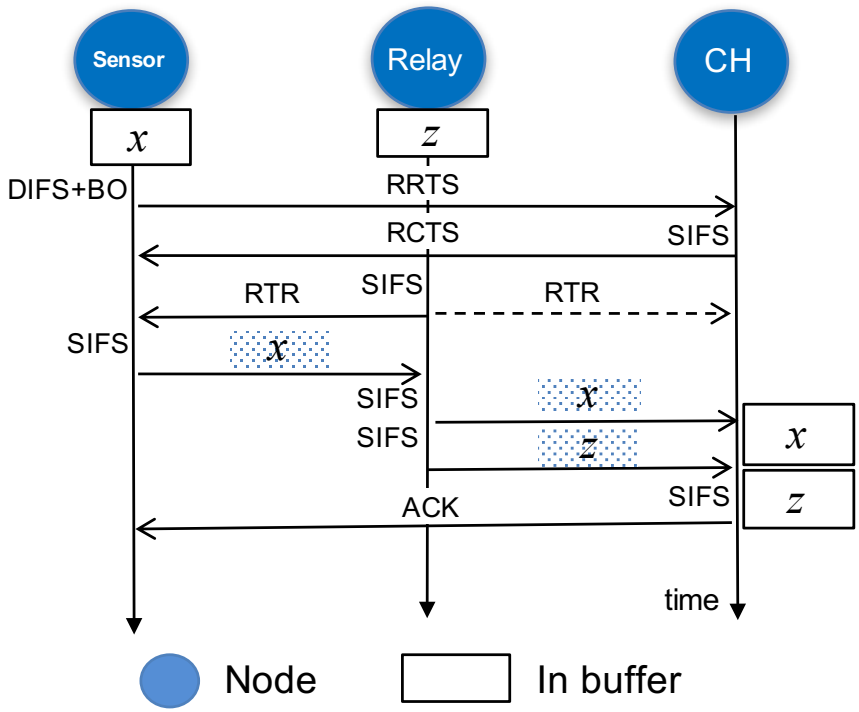


Figure 4.3: Timing operation of modified 2PSP protocol

The relay broadcasts a Ready-to-Relay (RTR) message. Firstly, the relay node determines a suitable pair of higher data rates based on the signal strength of the receiving RRTS and RCTS messages. Then it calculates the energy consumption and rate with one-hop transmission and with two-hop transmission. If two-hop transmission saves more energy, the relay broadcasts a Ready to Relay (RTR) message. This message contains the information about the selected rates that the sender should use when it sends the

data packets. If the sender cannot hear any response after SBI interval times out, the sender will transmit the data packet according to the standard Distributed Coordination Function (DCF) procedure. If the sender can correctly decode the RTR message, it will transmit its data packet with the new data rate defined in the RTR message. Then the relay will forward the data packet it received to the destination node.

Extra data transmission policy from relay node to CH is added after the 2PSP procedure of the modified protocol. The relay transmits its own data packet after forwarding the sensor's packet. After receiving the two packets from the relay, CH sends the ACK to sensor node. The advantages of allowing the extra data transmission are that the relay node can directly access the medium without competing with other nodes. It can also directly use the higher transmission rate defined by the RTR mechanism of 2PSP protocol. More data packets can be received at the destination with a fewer channel access competition and therefore, the whole DCF procedure benefits the low transmission delay.

4.5 Numerical Simulation

This section evaluates the performances of the proposed scheme by means of simulation results obtained from a simulation coded in MATLAB. The IEEE 802.11a hardware specifications are used for the simulation with the simulation parameters described in Table 4.1.

4.5.1 Simulation Scenarios and Parameters

The simulation is carried out in two parts. The first part is for the evaluation of modified 2PSP protocol in a cluster with 50 sensor nodes. For this scenario, nodes are randomly placed in a 2-D area with 100 x 100 m^2 size where the sink (base station) is placed at the centre of the grid. The assumption is that the network is divided into some clusters with its own cluster head each. Clustering a network can be performed as described in the first phase of LEACH. Firstly, nodes organise themselves into clusters and a node is elected as the cluster head within each cluster based on the probability distribution decision. If a node has been assigned as a cluster head, it sends advertisements to its neighbours. The neighbour nodes decide to which cluster they join based on the signal strength of these

messages. Random sources are selected and data flows from these sources are routed to the cluster head. The nodes communicate with the cluster head via the neighbour nodes in multihop fashion. The sensor data is sent to the cluster head by using the modified 2PSP in the cluster. The modified 2PSP protocol is used to find a path in low energy consumption and latency with the help of a relay in data transmission to the cluster head. The number of data flows in a cluster is varied from 10 flows to 50 by increasing 10 flows each time.

The second part of the simulation is to study the performance of network coding and decoding at each cluster head and at the sink respectively. The cluster head performs RLNC network coding on the received data. The cluster head transmits the coded data to the sink. These linear coded packets are transmitted to the sink until the sink can recover all the original data packets sent from the sensor nodes inside each cluster. In LEACH, the cluster head performs data aggregation and TDMA is used for medium access control inside the cluster. With a certain amount of energy, the total number of transmissions received at the sink and energy consumption are calculated. The assumption of the link quality between the cluster heads and the sink is set to be at least 0.5. This means 50% of packets can be lost during transmission to the sink. The performance of the transmission scheme with RLNC is compared with the transmission of LEACH.

4.5.2 Simulation Results and Discussion

In this section, the discussion on the performance of the proposed scheme is presented based on the results obtained from the simulation. For the transmissions from the sensor nodes to the related cluster heads, the modified 2PSP protocol is evaluated compared with the conventional CSMA/CA protocol in terms of throughput and energy consumption.

4.5.2.1 Comparison of Energy Consumption

As the Figure 4.4 depicts, the total energy consumption increases with the increase in the number of data flows inside the cluster. The energy consumption is the energy used for the transmission of control messages and the successfully transmitted data packets. The amount of energy consumption by the modified 2PSP protocol for 10 data flows is 4.2302 mJ and 20.913 mJ for 50 flows. The amount of energy saving due to the modified 2PSP

Table 4.1: Simulation parameters for evaluation of necoDG scheme

Parameter	**Value**
RRTS size	20 bytes
ACK size	14 bytes
RCTS, RTR size	15 bytes
MAC header	34 bytes
Transmission rate	6, 9, 12, 18, 24, 36, 48, 54 Mbps
MAC protocol	CSMA/CA, Modified 2PSP
Transmission power	100 mW
Payload size	1000 bytes

protocol is about 15.2% in average. A relay is selected based on the energy consumption and the rate by comparing two consumptions; one with direct transmission and another with two-hop transmission. This is different from the work of [46], where a candidate relay is selected only based on the achievable higher rate that the relay will apply in forwarding.

4.5.2.2 Comparison of Throughput

Figure 4.5 shows how throughput varies with the increase in the number of data flows from the sensors. Throughput is defined as the amount of payload data that can be received at the destination (here the cluster head) for all the unicast data flows transmitted by the sources (sensor nodes). The 2PSP outperforms the CSMA/CA protocol by up to 14.9% improvement in average. The throughput decreases as the number of data flows increase. The reason for this behaviour is because of the high latency when data flows increases. High latency means low throughput. When the number of flows in the network increases, there is high possibility of competing the media access, which leads to the delay and affects the throughput performance. The latency of 2PSP is 13% reduced compared to that of CSMA/CA.

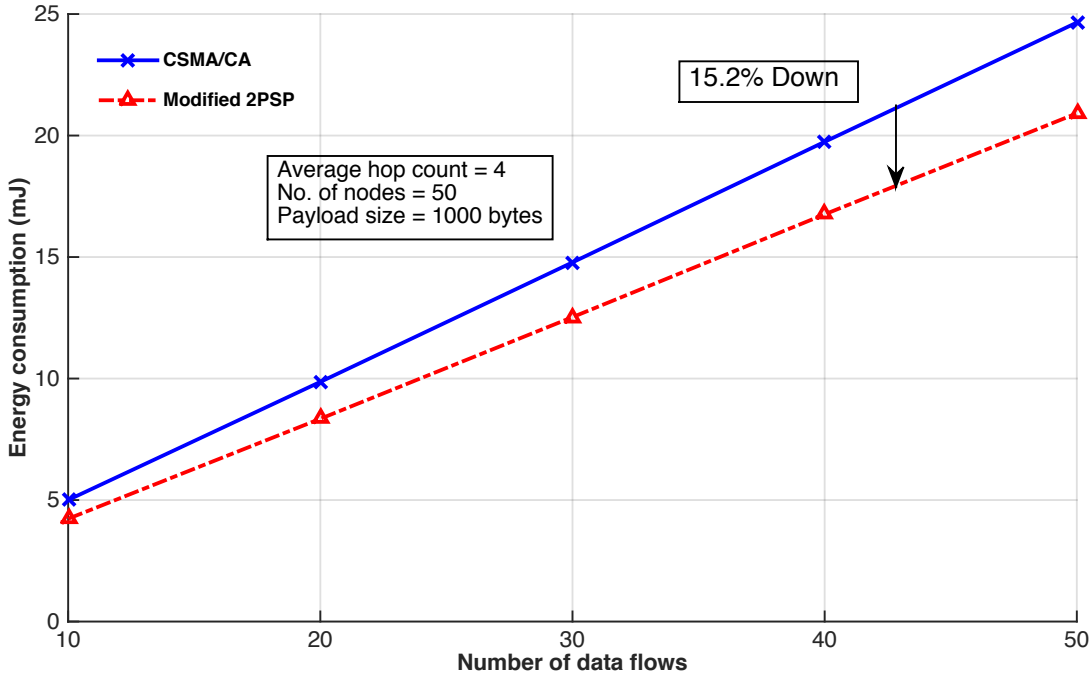


Figure 4.4: Energy consumption of 2PSP

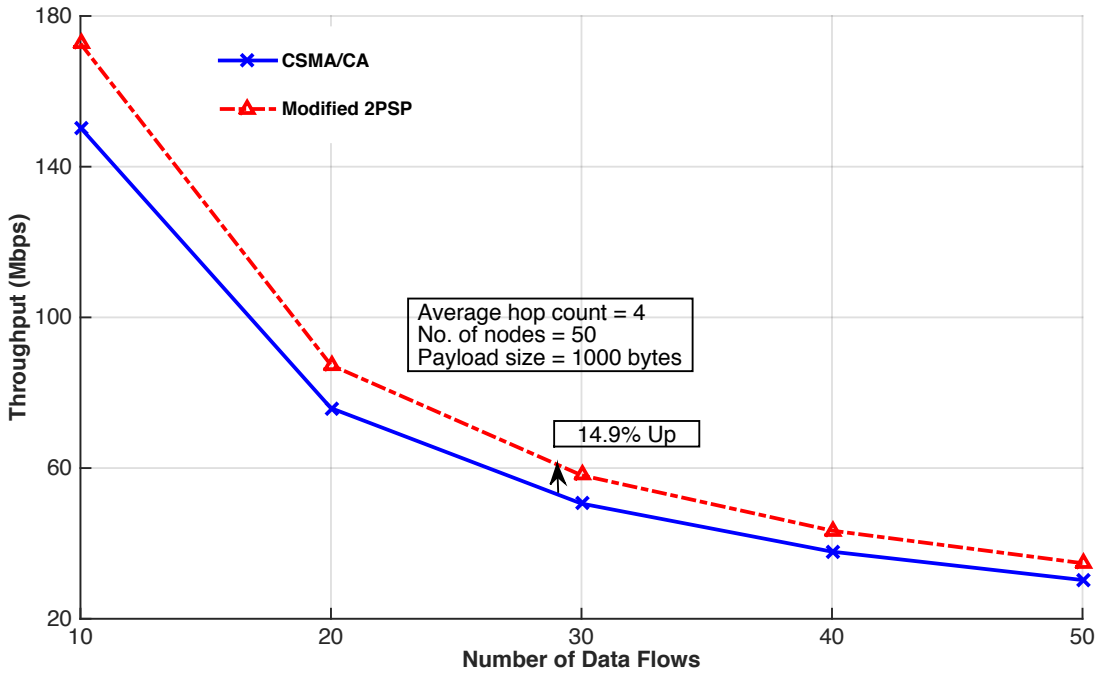


Figure 4.5: Throughput as function of flows increase

4.5.2.3 Comparison of Number of Transmissions

The performance of transmission with the RLNC network coding and LEACH are

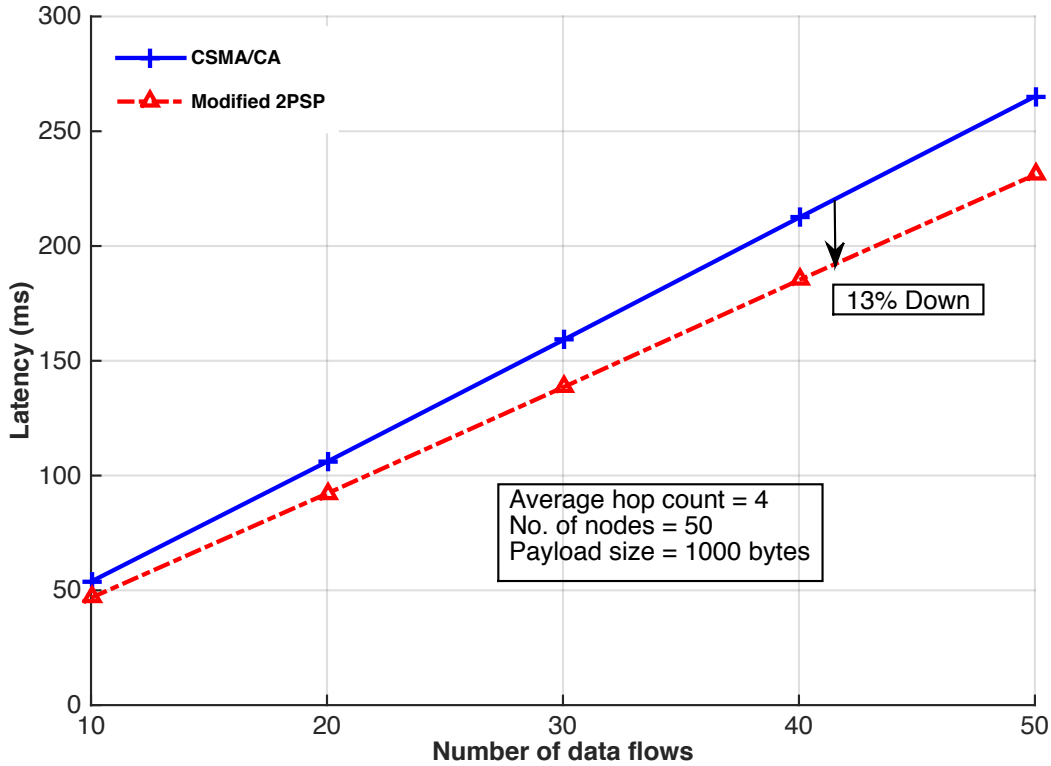


Figure 4.6: Latency as function of flows increase

compared for the cluster-based data gathering network architecture. Each cluster head node performs network coding function on the receiving data packets which include data from the own cluster and the overheard data from other cluster. Six data packets are encoded together in each generation. Transmission with RLNC coding scheme needed 25 transmissions with the link quality probability equal to 0.5 for the successful receiving of 6 original data packets. It costs 5.2 μJ for receiving 6 data packets. For the total 50 data flows from a cluster, the energy consumption is 0.26 mJ from the cluster head to the sink by the network coded transmission scheme and 20.913 mJ from 50 sensor nodes to the cluster head by the modified 2-PSP protocol. It means that it will totally cost 42.346 mJ for all the 100 data flows from the sensor nodes in two clusters while the number of transmissions by LEACH is 100 transmissions for energy 500 mJ of sensor nodes as described in their paper.

4.6 Summary

In this proposed necoDG scheme, RLNC is applied for data transmission between a sink and the cluster heads of each cluster inside the WSNs. A modified 2PSP protocol is also proposed for the data transmissions from sensor nodes to the corresponding cluster heads. The simulation results show that 2PSP can reduce the energy consumption in the data propagation process of sensors nodes up to 15.2%. The proposed transmission scheme with RLNC network coding can also reduce the number of transmission by encoding many packets together in each transmission. The 2PSP protocol can also be applied while the cluster heads transmit the network coded data packets to the sink for further increase the throughput and to reduce the energy consumption for the case of many cluster heads away from the sink.

Chapter 5

Balanced Cooperative Coding and Transmission (BCCT) Scheme

5.1 Introduction

The focus of this chapter is a scenario of cooperative data exchange which has the high prospect for future application. A balanced cooperative coding and transmission scheme called BCCT is proposed for this scenario, and it assigns suitable clients to transmit the encoded messages (linear combinations) of the received messages. The goals of the scheme are to minimise the total number of transmissions and to maintain the fairness among the participants in the cooperative group in order to save the limited energy resources until all members satisfy their needs. In addition, a transmission scheme with physical layer network coding called BCCT/PLNC is also designed for the same problem to further reduce the required transmission time slots and accomplish the data exchange process in short time. The performance of each scheme is validated in terms of fairness, the number of transmissions, and the completion time by testing with various simulation parameters and scenarios.

This chapter is organised as follows. Problem statement is described in Section 5.2, and the related work is discussed in Section 5.3. The scenario of the cooperative data sharing and the cooperative coding scheme is illustrated with example in Section 5.4. The detailed description of the proposed balanced coding scheme and the algorithm are presented in Section 5.5. Section 5.6 presents the introduction of physical layer network

coding (PLNC) into the data exchange problem. Simulation and discussions of the results are presented in section 5.7. Finally, Section 5.8 summarises the chapter.

5.2 Problem Statement

This chapter corresponds to the third problem presented in chapter 1. It is the end user needs for connectivity even in very crowded places such as stadiums, shopping malls, public events etc,. The invention of computing machines is also enhanced from the generation of supercomputers to the today's smartphones which already possess the processing and storage capacities of the earlier computers. Most people in modern society possess a smartphone. The conventional way of communication between two mobile wireless devices such as smartphones and laptops is through a base station which usually locates at a far distance because of the high expense of building towers for both cellular and WiMAX wireless base stations. An interesting problem scenario is the direct data sharing among the wireless client devices that are located in the vicinity of each other to reduce the amount of expensive communication to a far base station.

Due to the increase number of connected devices to the IP network and high data rate services such as video streaming and file downloads, the traffic volume is very high and leads to network overload.

Assuming that some nearby wireless client devices download a file from a common base station. Some devices may receive the whole file while some receive only a portion of it and lose other subset of the file because of the poor wireless links between the base station and them. In this case, the nearby wireless devices can utilise the advantage of short-distance communication links by creating a cooperative data sharing group among them via the built-in Bluetooth or Wi-Fi interfaces like in [8–11]. With this approach, devices will possess a faster and reliable short-range communication service to fulfil their requirements and no need to request to a far base station with slow connection again. There is no need to transmit to each user of the group again because all the n packets of the file are transmitted to the group. Consequently, the load on the base station will be reduced and the base station will be able to serve other users.

The users only need to collectively possess n packets within their group and want to cooperate to share the packets they have already received. The cooperative commu-

nication is predicted to become one of the major features for the emerging fourth and fifth generation wireless systems in [71, 72]. Cooperation between neighbouring nodes can provide many benefits such as energy saving of a device by means of sharing the load among them, bandwidth saving and low delay services because the expensive long-distance cellular links are free for another users. Moreover, the overall throughput of the network can be increased by exploiting the broadcast nature of the wireless medium. One important research problem is to maintain the fairness among the members of the cooperative group and to develop a mechanism that will handle it while also maintaining the minimum number of required transmissions.

5.3 Related Work

The problem of selecting an encoding scheme that will minimise the number of transmissions is referred to as index coding. The index coding problem in [25], where a central base station performs the transmissions of linear combined packets to make other clients satisfy their requirements. The difference is that, in cooperative approach, all of the clients involve in the transmission process. In this category, many existing research works such as [10, 73–75] mainly studied the optimum number of transmissions to reduce the complexity, overhead and delay until all clients ultimately recover the required packets. They formulate the problem into integer programming and proved that it is NP-hard.

With the application of linear network coding [26], the benefits of cooperation can be further expected for the data exchange problem because many devices can simultaneously gain from one linearly coded packet (linear combination). The assumption is that mobile devices are within the transmission range of each other and they can hear the broadcast transmission of others successfully. In this linear approach, the bits of each incoming packet is replaced by the symbols over the finite field $GF(2^n)$ and forms a vector of symbols. A client device creates linear combination of its packets and choose one combination to be transmitted that will benefit the other peers. The operations are performed over $GF(2^n)$. As other peer nodes in the cooperative group may already have some independent packets, they can recover a new one after performing Gaussian elimination of the received packets.

Rouayheb et al. in [10] formulate a lower and upper bound on the number of transmis-

sions needed to satisfy the requirements of all clients and they show that their algorithm performs closer to the lower bound by a numerical simulation. The authors choose a candidate client which possesses the maximum number of received packets as a transmitter for next round. If there is more than one candidate client, their scheme chooses the next transmitter randomly. Fairness has not been considered on the number of transmissions each client makes.

The fairness among the peers is considered in this chapter, and it is considered based on the number of transmissions from each peer. The scheme also maintains the minimum number of transmissions as a whole. The reason is that energy saving is important not only for the whole data exchange process but also for each individual participant. If a client which possesses an independent packet runs out of energy quickly, other clients will not satisfy their needs and the data exchange process will not accomplish, by assuming that client devices collectively possess all messages from the base station for a certain data file. An algorithm is proposed for these purposes, distributing the number of transmissions among the peers while maintaining the total number of transmissions to be minimum. In this algorithm, each client changes the role of transmitter based on the reception knowledge they keep in information table in each round of data transmission.

5.4 System Model

The scenario of a cooperative data sharing is shown in Figure 5.1. In this scenario, some nearby mobile devices download a file from a common base station which may be located at a long distance. After the data messages of the file are transmitted by the base station, some client devices may only receive some messages but all the nearby nodes collectively receive all the messages. The problem is to satisfy the required messages of all peers in the cooperative group by sharing each other via the local wireless connections, needing no request to the far base station. This problem can be formulated as a cooperative linear network coding scheme. For the basic understanding of a cooperative linear network coding scheme, a scenario with four clients and total six data messages is described in the information table 5.1 to illustrate the example in Figure 5.1 and, to explain how it works. At the initial state, clients c_1 and c_3 have received five messages out of six and are the clients that have maximum number in the cooperative group. Client c_2 has received three

messages and therefore, it is the client with the least number of messages while client c_4 receives four. The linear network coding scheme calculates an encoded message which is a combination of multiple messages from a client and transmits the encoded message. Many other client devices with the requirement of different messages can benefit from the same encoded message.

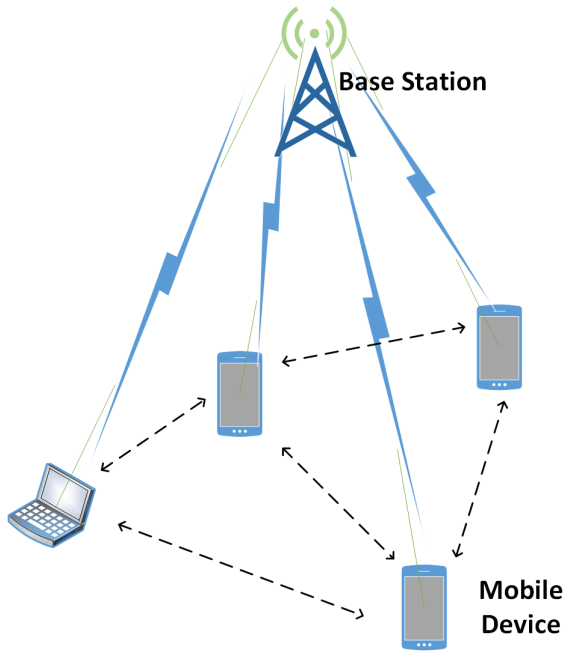


Figure 5.1: Cooperative data sharing among the nearby mobile clients after downloading a file from a far base station.

In the example, the data exchange among four clients accomplishes with three coded messages transmissions; two transmissions: $x_1+x_2+x_3+x_4+x_6$ and $x_1+x_2+x_3+x_4+x_5+x_6$ from the maximum client, c_3 and one transmission: $x_1+x_2+x_3+x_5+x_6$ from the another maximum client, c_1. However, it should be more desirable that the coding scheme can distribute the transmissions among as many clients as possible in order to maintain fairness among them. As the clients will gradually increase the number of received messages iteration by iteration, the dependency between the encoded message and the existing message becomes higher. In this situation, even choosing a client with the maximum number of messages will not provide significant improvement. Giving higher priority to a client which has less participation in the data exchange process will share the workload of previous transmitters. As a consequence, those nodes in the cooperative group can save their energy and computing resources. This is the main idea provided in the proposed

balanced coding scheme to keep fairness among as many client devices as possible.

5.5 Proposed Balanced Cooperative Coding and Transmission (BCCT) Scheme

Many important components comprise the proposed balanced coding scheme and they are depicted in a conceptual block diagram by linking many other components of data link layer as shown in Figure 5.2. The purpose of this scheme is to accomplish the data sharing process in the group with less transmission time and to keep fairness among the participants in the group. The group formation management component includes the techniques based on creating clusters or dividing the network according the zones or regions. The proposed BCCT scheme is not intended for the group formation process, but it is for the data sharing process within the group. Buffers are needed for the incoming and outing messages; specific functions such as coding and decoding operations, the selection of a best combination; selecting a transmitter based on the knowledge of receiving messages at the clients and accessing a channel to broadcast a coded message all together comprise in the BCCT scheme.

When the number of messages involved in one generation of data exchange process is high, the computation for the coding and decoding of the coding scheme can become burden for a mobile node. However, with the rapid development in smartphone technologies, this burden will soon be covered. Padding of zeros is considered for the addition of two messages which have different lengths. In consideration of these problems, it is assumed that the wireless nodes involved in the data exchange have high processing and memory resources for computing and buffer requirements.

Buffer Management at Each Node

Two buffers are assigned in each client device: one for the incoming messages and another for the outgoing coded messages. The recovered original messages will be stored in the permanent storage. At the start of a cooperative data sharing process, one or more peer devices will have received all or some of the messages from the base station. Therefore, the size of the incoming buffer is equal to the number of messages n and that of the outgoing

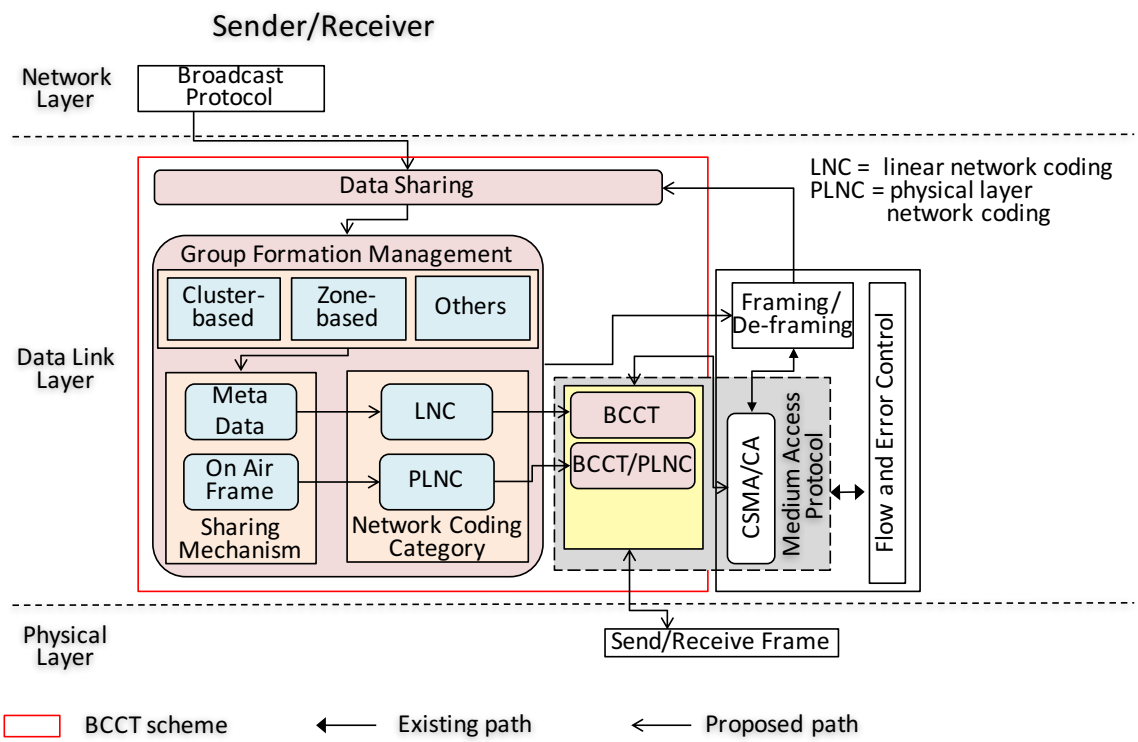


Figure 5.2: Block diagram of Balanced Cooperative Coding and Transmission (BCCT) Scheme

buffer is equal to size of the block based on how many messages are coded together.

Reception Information Table for Next Transmitter Selection

Maintaining fairness among the participants of the data sharing group is one of the goals of BCCT scheme. To achieve this goal, an algorithm is proposed to select the transmitter for each iteration based on the number of messages each client node possesses, usually the client with maximum messages. Moreover, nodes should also take into account the number of transmissions they made in order to maintain the fairness among them as explained in Section 5.4. The approach is to keep an information table like Table 5.1 in every node. At the start of data sharing process after receiving messages from the base station, all clients should broadcast reception information to announce their packet receiving status. Then a client node can use this knowledge for deciding a suitable transmitter for next round. This can be described as a sharing mechanism in the block diagram of BCCT scheme.

Table 5.1: Example of reception information table: index of received packets and number of transmissions from each client station

Clients	**Data messages**						**Transmissions**
	x_1	x_2	x_3	x_4	x_5	x_6	
c_1	1	1	1	0	1	1	1
c_2	1	1	0	0	0	1	0
c_3	1	1	1	1	0	1	2
c_4	1	1	0	1	0	1	0

Coding and Decoding

Coding and decoding component is one of the main functions of BCCT scheme for performing the coding operation to find a suitable combination of messages to be transmitted and for decoding the received combination with the existing messages. Then the successful client as a transmitter calculates the linear combinations of messages from the incoming buffer by using the encoding vectors and stores the coded messages in the outgoing buffer before being sent. After each round of coded message transmission, every node performs decoding operation of the incoming coded message and the existing received messages. Then the clients update their information table and the reception reports are broadcast to decide who should take turn for transmission in next round until all the clients satisfy their requirements. This is possible due to the fact that all nodes exist within the transmission range and every node can hear the transmission from each other.

Medium Access Control (MAC) Mechanism

After receiving some initial data messages, each client calculates and broadcasts their packet reception information (metadata) in short control messages. Then they will know who is supposed to start transmitting. The clients with maximum number of messages will basically transmit in every iteration. After each iteration, they update their received messages. The transmitter can also predict their reception because it calculates during choosing best combination. However, at later iteration, to keep balance, the transmitter is selected based on the fewer number of transmission in the case they have received same

number of messages. In order to provide a high decoding probability, each client station needs to know which messages other members of the group have received to determine the optimal coding decision. To provide this reception information, each station includes a reception report in the data frame header. The reception report contains the list of currently buffered messages. Zhang et al. in [76] use a method to turn a client device in master role (transmitter) to slave role (receiver) by sending the access request message for the Piconet-based distributed cooperative approach, which is similar to the data exchange problem here. In the proposed algorithm, the winner for transmission will broadcast the coded messages and other clients refrain from transmitting for the specified duration.

5.5.1 Algorithm for Balanced Coding Scheme

At each iteration of the algorithm, one of the clients broadcasts a linear combination of messages in its incoming buffer which are from the set $X = \{x_1, \cdots, x_n\}$ of n messages. A coded message x is denoted by $C_x \in GF(2^n)$, the corresponding vector of linear coefficient, i.e., $x = C_x.(x_1, \cdots, x_n)^T$. Y_i is denoted as the subspace spanned by the vector corresponding to the linear combinations available at client c_i. At the beginning of the algorithm, Y_i is equal to the subspace spanned by vectors that correspond to the messages in $X_i \subseteq X$, i.e., $Y_i = \langle\{C_x | x \in X_i\}\rangle$. The goal of the algorithm is to simultaneously increase the dimension of the subspaces $Y_i, i = 1, \cdots, k$, for as many clients as possible. At each iteration, the algorithm identifies a client $c_i \in C$ whose subspace Y_i is of maximum dimension. Then, client c_i selects a vector $b \in Y_i$ in a way that will increase the dimension of Y_j for each client $c_j \neq c_i$, and transmits the corresponding packet $b.(x_1, \cdots, x_n)^T$. At some iteration, the subspaces associated with many clients may become identical. This group of clients are merged into a single client with the same subspace.

5.5.1.1 Algorithm for Best Combination Selection

In the linear data exchange scheme, each client should calculate the receiving information to choose a suitable combination for the transmission. In principle, the client with the maximum number of packets possesses the best combination at earlier stages of the process as the dependency between the combination and the received packets at other peers is low. Selecting the best combination that can increase the dimension of the subspaces of as

Algorithm 1 Algorithm for BCCT

```
for i = 1 to k do
    Y_i = ⟨{C_x | x ∈ X_i}⟩
end for
while there is a client i with dim Y_i < n do
    while ∃c_i, c_j ∈ C, i ≠ j, such that Y_i = Y_j do
        C = C \ {c_i}
    end while
    Find a client c_i with a subspace Y_i of maximum dimension
    if there is only one c_i then
        Select a vector b_i ∈ Y_i such that b_i ∉ Y_j for each i ≠ j
        Let client c_i broadcast packet x = b.(x_1, ··· , x_n)^T .
        Store c_i as transmitter in the table
    else
        Find a client c_j with a subspace Y_j of maximum dimension
        Select the client that has fewer previous transmissions
        if c_i is chosen then
            Select a vector b_i ∈ Y_i such that b_i ∉ Y_j for each i ≠ j
        else
            Select a vector b_j ∈ Y_j such that b_j ∉ Y_i for each j ≠ i
        end if
    end if
    Let client c_i or c_j broadcast packet x = b.(x_1, ··· , x_n)^T .
    for l = 1 to k do
        Y_i ← Y_i + ⟨{b}⟩
    end for
end while
```

many clients as possible is one of the major goals of the algorithm. One major challenge for this job is the computational time required to find it for some cases depending on the distribution of the initial receiving messages. To overcome this challenge, a variable is introduced in the proposed algorithm, it will control the number of clients that should

Algorithm 2 Algorithm for selecting best combination

if (number of other clients with $rank = max\_rank$) $\geq$ (number of other clients - 1) **then**

Set control variable = 1

Calculate best combination

else if (number of other clients with $rank \geq max\_rank - 1$) $\geq$ half of number of other clients **then**

Set control variable = 2

Calculate best combination

else

Set control variable = number of other clients

Calculate best combination

end if

increase their ranks of the subspaces. The value of this variable is derived by studying the relationship between the ranks of the subspace of each client after running the simulation many times. For those cases, the chosen combination might not be an optimum one. The condition and the value of control variable for picking the best combination is described in Algorithm 2.

5.6 Balanced Coding and Transmission Scheme with PLNC (BCCT/PLNC)

The concept of physical layer network coding (PLNC) is applied to the data exchange problem to further reduce the number of transmissions and accomplish the data exchange process quickly. The basic idea of PLNC is to exploit the mixing of signals that occurs naturally when electromagnetic (EM) waves are superimposed on one another [38]. The simultaneous transmissions by several transmitters result in the reception of a weighted sum of the signals at a receiver. The advantage is that the number of required time slots can be reduced because transmission from two clients is allowed in the same time slot. The relay is simply an intermediate node that receives the addition of combinations transmitted from the client nodes and broadcasts the coded message back to the clients

including the transmitters. In this scheme, relay is one of the other nodes in the group.

By taking advantage of this combination, the MAC is scheduled to allow transmission from two clients in one round of the algorithm. Their addition will be received by the other members of the group. This weighted sum is a form of network coding operation by itself. Therefore, even if the receiver forwards the receiving combination, the two transmitters can themselves benefit from the addition of coded messages. It is the advantage of this transmission scheme. Two stations with maximum rank are decided as transmitter by exchanging short control messages. In this decision, their number of transmission is also taken into account to maintain fairness. In the control message, the time to start transmission is also defined. The second station replies with short control message. Other stations that overhear the message will stop channel access until the transmissions from two stations finish and wait for receiving. After receiving the PLNC-coded message, they perform decoding and updating their buffer. Then a station starts transmission of control message together with the updated reception information if it has the maximum rank after decoding. This process is repeated many times until they notice that the reception information is identical to the previous iteration. After that, they start to change the transmission mode to the one without using PLNC technique.

The drawback of this approach is that the mixed signal sometimes results in no increment in the rank of the subspace of other clients. This is because when the client devices are gradually increasing their received messages, the dependency of their linear combinations are also high. In such conditions, transmission in PLNC has no advantage. This will be true especially for the later iteration of the transmission process. To avoid this drawback, BCCT/PLNC will apply more transmission in PLNC in the earlier stages of process and will stop transmission in such mode at the later iteration. Transmitting only one combination will be continued. The significance of PLNC over the LNC is that the transmitted combination can carry more information and more number of client stations can benefit especially at the earlier iteration of the data exchange process. Therefore, the required transmission time slot can be reduced.

5.6.1 Algorithm for BCCT/PLNC

The idea of this algorithm is to allow transmitting the addition of two linear combinations from two clients. When the messages to be combined do not have the same length, the shorter ones are padded with trailing 0s to make their lengths identical. The resulting encoded messages (linear combinations) have same length as original ones. First, the algorithm chooses a linear combination from the client with maximum number of messages. Then it finds another linear combination from the second client. These two combinations are transmitted simultaneously and allowed to add in the air naturally when electromagnetic (EM) waves are superimposed on one another. Relay node deals with the mapping of the mixed signal to the desired network-coded signal $S_R = S_1 \oplus S_2$. Then, it broadcasts S_R to other clients in the second time slot. In the algorithm, two combinations are added first and then transmitted their addition. If there is more than one client with the maximum number of messages, the two transmitters are carefully designated based on the information of their previous transmissions as in the balanced coding scheme.

5.7 Numerical Simulation

A computer-based simulation is coded in MATLAB to evaluate the performance of the proposed algorithms and transmission in PLNC scheme. The finite elements from $GF(2)$ are used for indexing and linear network coding of each received and lost message of each client for less complexity. The various aspects and performance metrics are studied in the simulation to evaluate the performance of the proposed schemes.

5.7.1 Simulation Scenarios and Parameters

Firstly, a comparison of transmission in balanced scheme and the transmission without balanced scheme (i.e., nominating the next transmitter randomly) is studied. Then, the impact of the number of messages and client nodes involved in the cooperative data sharing on the fairness of the system is also investigated. The impact of initial packet receiving probability P_{init} on the system performance is also analysed by different probability values with total 10 messages.

In the second scenario of simulation, the number of time slots required by the BCCT/PLNC

Algorithm 3 Algorithm for BCCT/PLNC

```
for i = 1 to k do
    Y_i = ⟨{C_x | x ∈ X_i}⟩
end for
while there is a client i with dim Y_i < n do
    while ∃c_i, c_j ∈ Ci ≠ j, such that Y_i = Y_j do
        C = C \ {c_i}
    end while
    Find a client c_i with a subspace Y_i of maximum dimension
    if there is only one c_i then
        Find a client c_j with a subspace Y_j of smaller maximum dimension than Y_i
        Select a vector b_i ∈ Y_i such that b_i ∉ Y_j for each i ≠ j
        Select a vector b_j ∈ Y_j such that b_j ∉ Y_i for each j ≠ i
        Add two vectors, b = b_i + b_j
        Let client c_i and c_j broadcast packet x = b. (x_1, ··· , x_n)^T .
    else
        Find a client c_i with a subspace Y_i of maximum dimension
        if dimension dim Y_i = number of packets then
            Select a vector b ∈ Y_i such that b ∉ Y_j for each i ≠ j
        else
            Select a vector b_i ∈ Y_i such that b_i ∉ Y_j for each i ≠ j
            Select a vector b_j ∈ Y_j such that b_j ∉ Y_i for each j ≠ i
            Add two vectors, b = b_i + b_j
        end if
        Let client c_i and c_j broadcast packet x = b. (x_1, ··· , x_n)^T .
    end if
    for l = 1 to k do
        Y_i ← Y_i + ⟨{b}⟩
    end for
end while
```

scheme is investigated. This is to find out how PLNC can help the data sharing process accomplish earlier than the scheme without PLNC. The total number of time slots required is expected to be fewer. The performance of the whole coding scheme is evaluated with respect to the increasing number of messages and client devices in terms of scalability and fairness respectively. The parameters for the simulation are described in Table 5.2.

Table 5.2: Simulation parameters for evaluation of BCCT scheme

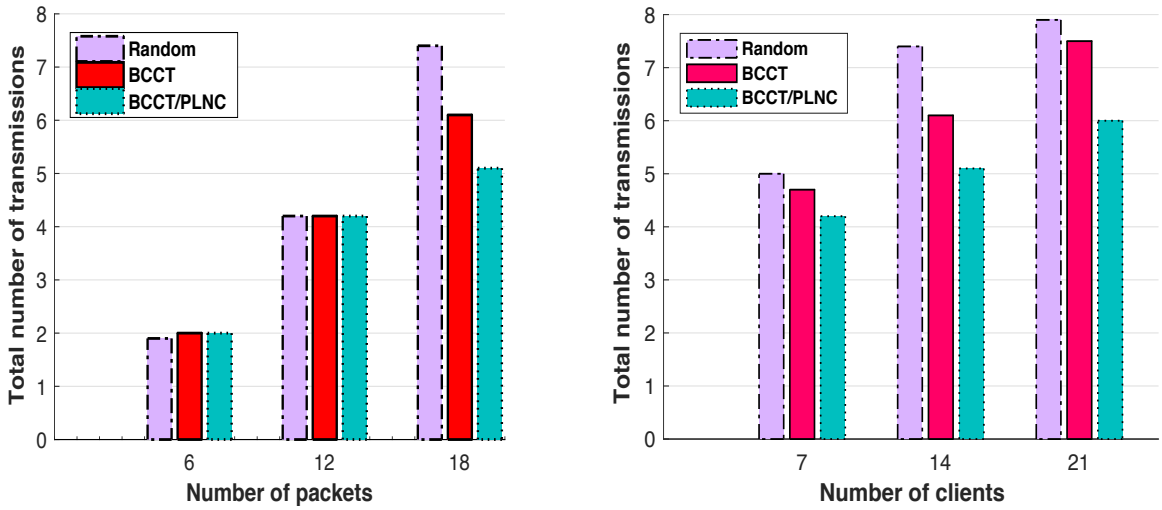
Parameter	**Value**
Hardware specification	IEEE 802.11a OFDM
Antenna type	Omni Antenna
Simulation environment	MATLAB 2014b
Initial packet receiving probability P_{init}	0.9, 0.7, 0.6, 0.5, 0.3
Number of messages	4 to 21
Number of clients	4 to 18
Finite field elements	$GF(2)$
Number of experiments	30

5.7.2 Simulation Results and Discussion

This section presents the performance of the proposed balanced coding and transmission scheme with PLNC based on the results obtained from the simulation. The discussions are mainly on the number of transmissions from each client in BCCT scheme, the number of clients whose ranks increase in each round of process in BCCT/PLNC scheme and the total number of transmissions and time slots applied for the entire data exchange process. The results for each scenario are the values averaged over the 50 experiments for 5 to 21 clients downloading 6 to 18 messages from a base station with the initial receiving probability of 0.3 to 0.9.

5.7.2.1 Comparison of Total Number of Transmissions

The total number of transmissions required by the proposed schemes and the original scheme for the completion of one data exchange process is shown in Figure 5.3. The proposed schemes need fewer number of transmissions compared to the original one with random selection. The scalability of the system is studied with the number of messages and client stations increase. The BCCT/PLNC scheme requires fewer number of transmissions than the BCCT scheme for the higher number of messages. Both the BCCT



(a) Different number of messages with 14 clients

(b) Different number of clients with 18 messages

Figure 5.3: Total number of transmissions of three different schemes with Increasing number of messages and Increasing number of clients.

and BCCT/PLNC schemes performs well as depicted in Figure 5.4 because the number of transmissions is within the upper and lower bounds as defined in the reference paper [10].

In Figure 5.5, the impact of the initial packet receiving probability P_{init} to the total number of required transmissions is shown. The results are for three values of P_{init} with 0.3, 0.5 and 0.6. Obviously, with $P_{init} = 0.3$, the number of initial receiving messages is a few and more number of transmissions are needed than that of $P_{init} = 0.5$ and $P_{init} = 0.6$.

Figure 5.6 shows that BCCT scheme can significantly distribute the transmissions among the participants in cooperative data exchange. In Figure 5.6(a), client c_4 is the

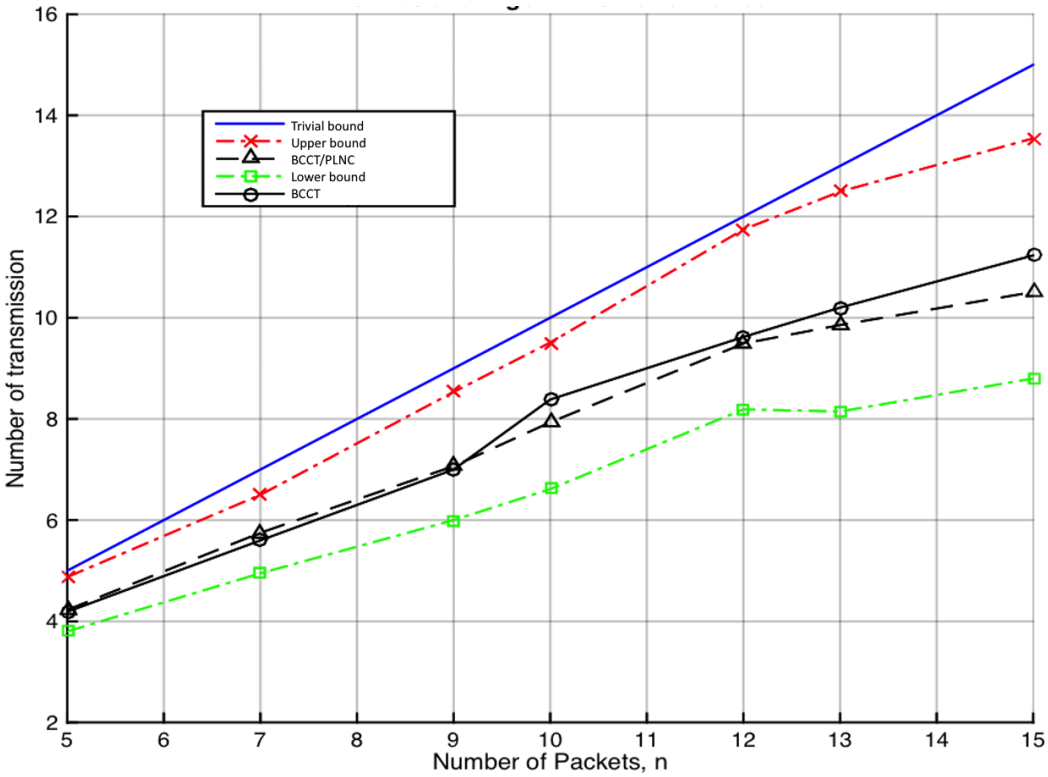
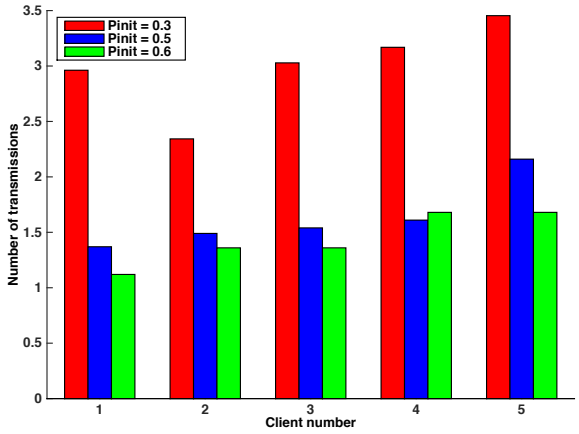


Figure 5.4: Performance of the algorithms with the increased number of messages. The client number is 5.

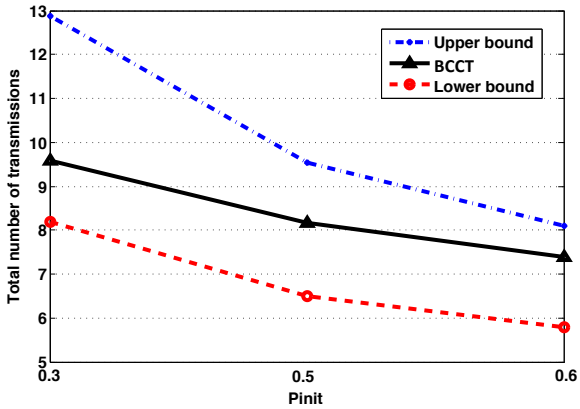
only client with the maximum number of transmissions. This condition disappears in the proposed balanced coding scheme. The algorithm tries to balance the transmissions of the clients. This can be easily seen in the transmissions of clients c_2 and c_3; and clients c_4 and c_5 compared to the original scheme without balancing capability. But the client c_1 shows no significant change in the balanced scheme. This is because client c_1 is the client with the minimum number of messages initially received and it has the least chance for participation in the data exchange process. Client c_1 is a receiver in most of the iterations of data transmission. The results also show that the average number of transmissions in the proposed balanced coding scheme is even fewer than the original scheme without balancing capability. Without balancing scheme, one or more clients have high number of transmissions. The proposed balanced coding scheme distributes the number of transmissions among the client devices.

5.7.2.2 Fairness on the Number of Transmissions

With reference to the previous figure, the number of transmissions from each client



(a) Transmission from each client



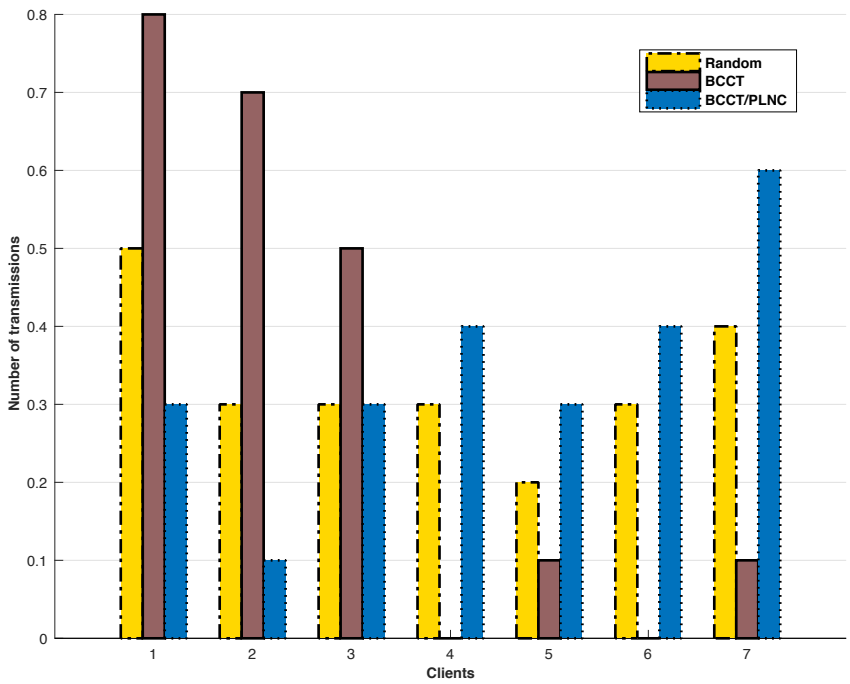
(b) Total number of transmissions

Figure 5.5: Impact of initial packet receiving probability P_{init} for 5 client devices and 10 messages

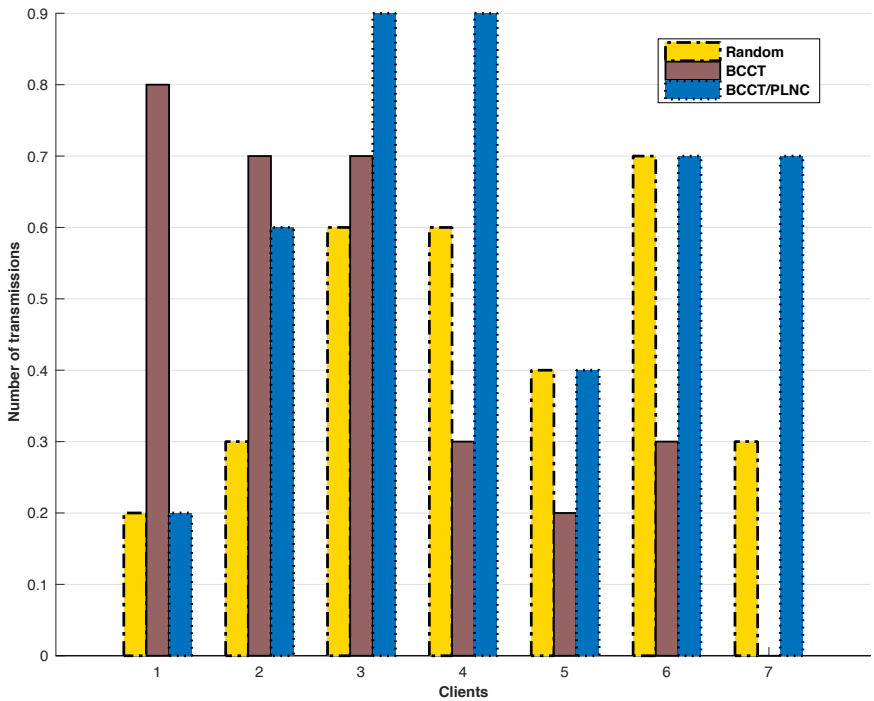
varies with the increase members in the group. In this section, the fairness of the three schemes is discussed for different number of clients and messages in Figure 5.7. The fairness of each scheme is calculated based on the number of transmissions from each client station by the Jain's fairness index,

$$\frac{(\sum x_i)^2}{n \sum x_i^2}, \tag{5.1}$$

where n is the number of stations participated in the data sharing group and i represents each participant. x is the number of transmissions from station i. The value of 1 means the best fairness index. The proposed balanced coding and transmission (BCCT) scheme achieves the highest fairness index among the three schemes for all different number of

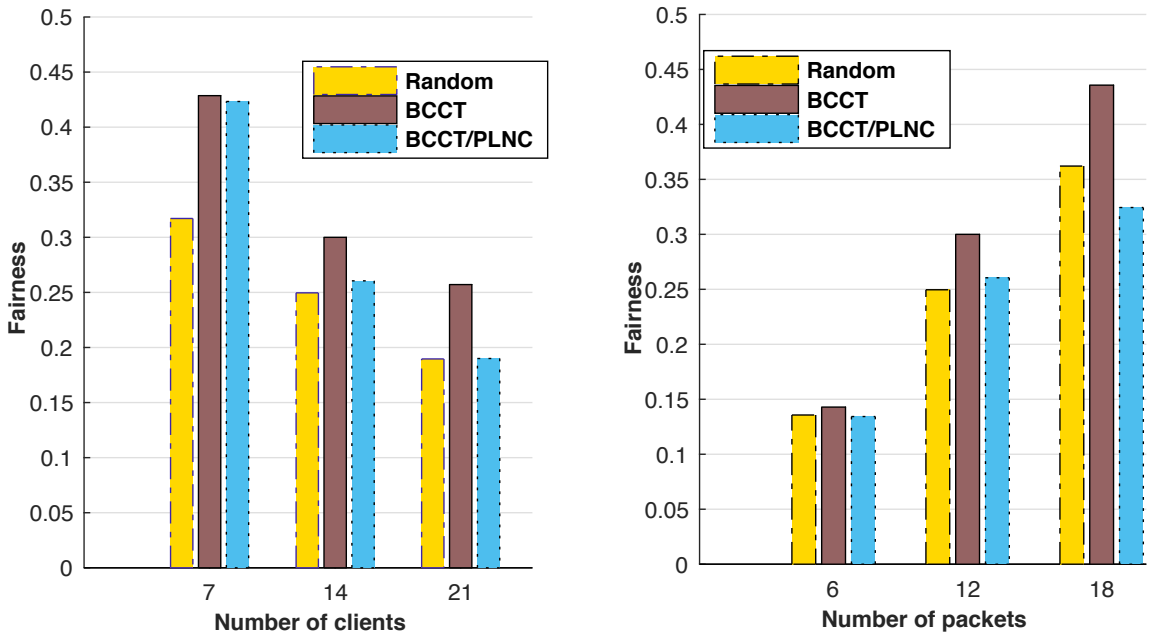


(a) With 7 clients and 6 messages



(b) With 7 clients and 12 messages

Figure 5.6: Number of transmissions made from each client for 6 messages and 12 messages.



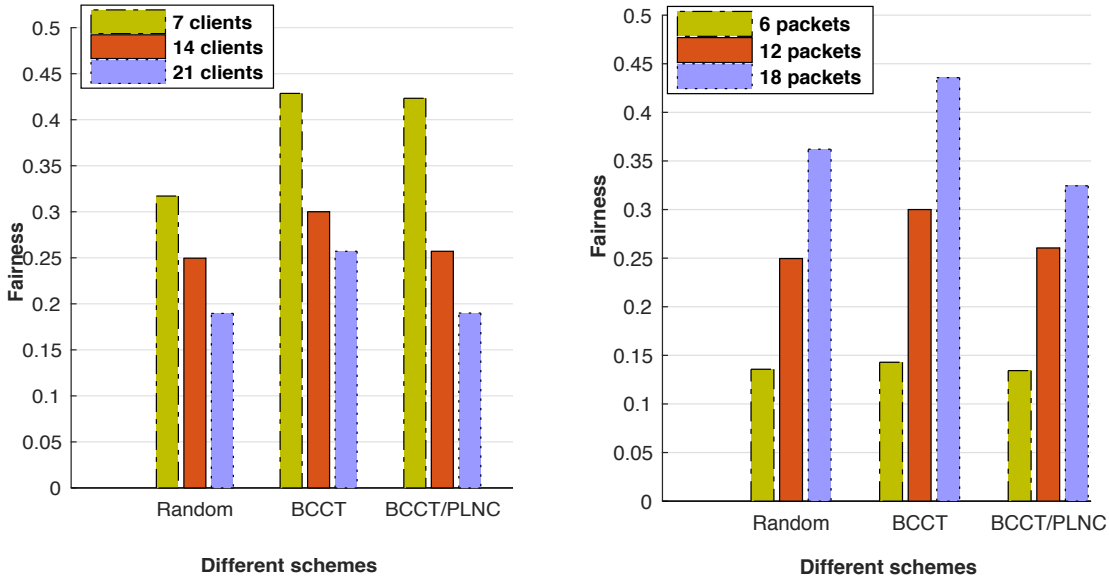
(a) Fairness versus different number of clients with 12 messages

(b) Fairness versus different number of messages with 14 clients

Figure 5.7: Fairness of proposed schemes versus that of random scheme for different number of clients and messages. Fairness of three schemes with respect to the number of clients and number of messages

clients and packets. The BCCT/PLNC scheme sometimes shows the lower index than the Random scheme. This is because the BCCT/PLNC scheme gives more priority for providing the required messages as quickly as possible. However, it still can provide the fairness index as high as that of the random scheme most of the times.

The impact of the increasing number of clients and messages over the fairness is also investigated. Figure 5.8 shows that fairness is affected by the increasing number of clients in the cooperative group and the number of messages for each downloading session. Fairness indexes of all the three schemes decline with the increase in the number of clients. This is due to the fact that there is a high chance of less participation from many clients. Their requirements are satisfied before many of them get the chance to transmit. After many transmissions from the clients with higher number of received messages, all clients satisfy their needs and no chance for participation any more. Reversely, the fairness value increases with the increase number of messages because more transmissions are needed



(a) Fairness versus increasing number of clients with 12 messages

(b) Fairness versus increasing number of messages with 14 clients

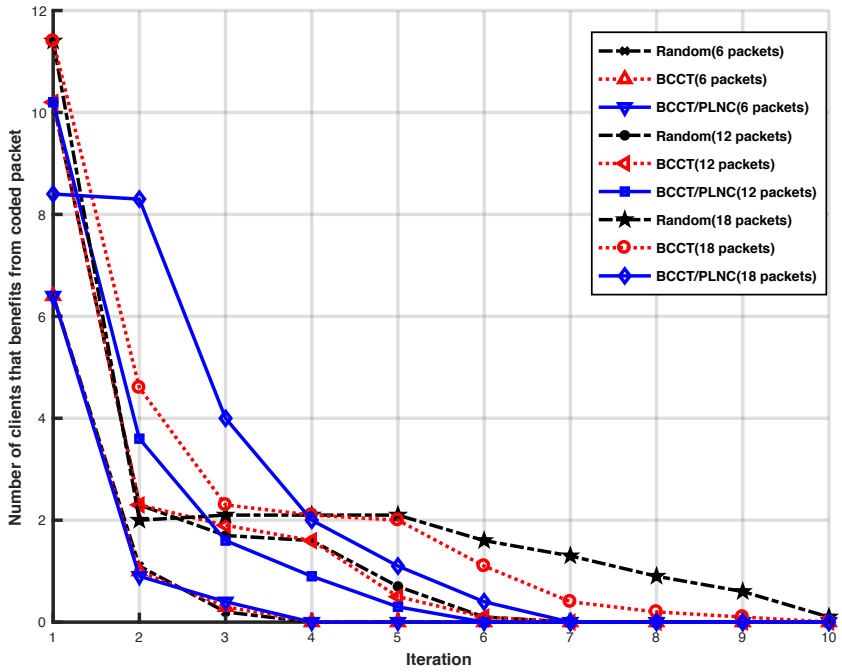
Figure 5.8: Impact of the increasing number of clients and messages over fairness.

and more clients have chance to participate.

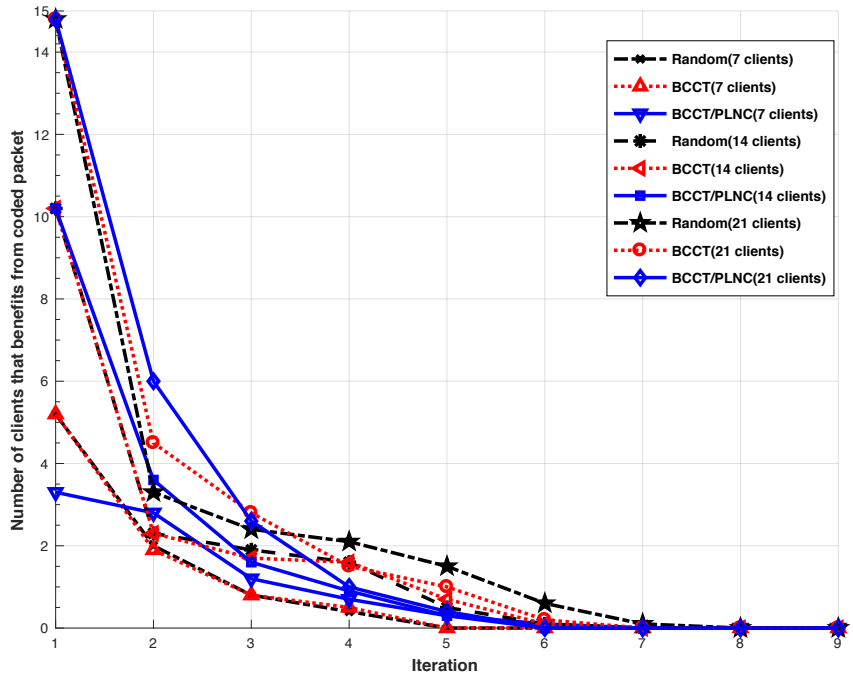
The fairness of the BCCT/PLNC is not as good as that of BCCT. This is because the main purpose of BCCT/PLNC scheme is to reduce the required time slot and number of transmissions. The selection of transmitters is based on the higher priority given to improve the coding operation. At the later iteration, the random node participates in forwarding the coded messages. From the standpoint of clients, they still need to involve in the simultaneous transmission until the algorithm notices the redundant transmissions.

5.7.2.3 Clients that Benefited in Each Round due to PLNC

The number of benefited clients in each iteration of the data sharing process is depicted in Figure 5.9. This value is high for BCCT/PLNC scheme in earlier iteration, meaning that many client devices can recover their lost messages from the PLNC-coded messages. The graph gradually declines with some fluctuation as the percentage of requirements of the clients decreases in later iteration. The graphs of BCCT and random scheme are especially high in the middle of the process for both cases and at the later iterations of the process for the high number of messages. This behaviour shows that the PLNC



(a) Number of clients that benefits among total 14 clients for different number of messages



(b) Number of clients that benefits for the total 12 messages

Figure 5.9: The number of clients that benefited in each iteration of the data sharing process.

transmission scheme increases the ranks of the subspaces of many receivers as soon as the transmissions are started. BCCT/PLNC can help more clients to benefit from the coded packet and reduce their requirements at the later iterations than BCCT and the random scheme.

This is different from the BCCT and the random scheme, where a fewer number of receivers can benefit from the coded packet in each iteration. At time slot 1, the number of benefited receivers is high for nearly all schemes because the transmitted combination is innovative for many receivers. Suddenly this number declines in second time slot as many stations recover their lost messages in the previous iteration. BCCT/PLNC scheme benefits more clients quickly while the other schemes can benefit gradually. Therefore, BCCT/PLNC scheme can satisfy the needs of the many users in the group within a few transmission time slots. With the transmission with PLNC, most of the clients devices increase their ranks in the earlier round of process. This helps the data exchange process to accomplish quickly by satisfying the requirements of participants and also ensuring that all the devices in the cooperative group recover their lost messages. Therefore, bandwidth and energy resources are saved.

5.8 Summary

This chapter presents a balanced linear coding scheme and a transmission scheme with physical layer network coding for the data exchange problem, with the purpose of maintaining the fairness among the client devices, and for accomplishing the process in minimum number of transmissions. The schemes create an information table that keeps the packet reception information, and the number of transmissions from each client in each round of data exchange process. This knowledge is important to maintain fairness among the clients to ensure that a certain client does not run out of energy and leave the group. With the proposed schemes, the total number of transmissions decreases while distributing the work load among the clients. Moreover, by allowing two clients to simultaneously transmit their linear combinations by the physical layer network coding, the results show that clients receive their required messages mostly in the earlier iteration of the process. Consequently, it satisfies the requirement of participants as quickly as possible and leads to the quick completion of data exchange process. The number of transmissions is also

within the lower and upper bounds, and the algorithms maintain the scalability for the case of increasing number of packets.

Chapter 6

Conclusions and Future Work

6.1 Summary and Contributions

This dissertation has developed an efficient network coding based data transfer framework called E-neco framework. The purpose of the framework is to provide the efficient data transmission, data collection and data sharing in multihop communication scenarios. This work has based on the network coding techniques in the literature and looked for the way of how these techniques could be utilised to solve the problems of future wireless networks efficiently. Efficiency is defined as accomplishing tasks with less amount of resource usage such as bandwidth, energy and latency. The scenarios considered in this research have high popularity in the current situation and the near future. Each category of data transfer is focused by proposing a scheme to achieve the desired goals to provide the requirements of future wireless network: energy-efficient high-rate communication, low latency and high reliable communication, and quick download access in a fewer number of transmissions, respectively. Each work is recalled briefly as below.

In Chapter 3, a network coding-aware medium access control (necoMAC) scheme has been designed to achieve the higher throughput. The protocols in this scheme utilise the chain and triangle topologies as the golden resources for the network coding opportunity. The topology management mechanism is incorporated in the MAC protocol. The essence of this design is the combination of the property of network coding: reduction in the number of transmissions and the multi-rate multihop transmission capability of NCA-2PSP protocol. This scheme provides higher throughput and faster transmission from

node to node at a link-level transmission. As the network coded transmission opportunity is reserved for the relay node, it immediately can access the medium. This improves the performance of the scheme. The overall performance of the network is improved by 30% due to the proposed scheme. This work gained the high data rate per certain energy usage.

In Chapter 4, an energy-efficient network coding based data gathering scheme called necoDG for data gathering applications has been proposed. The modified 2PSP MAC protocol is used for the data transmissions from sensor nodes to the corresponding cluster head (CH). Some new rules are applied at the relay node to choose a path that costs less energy consumption. Random linear network coding (RLNC) is used at the aggregator node, CH to transmit coded packets from cluster heads to the base station. The network coded transmissions are suitable to reduce the control messages in case of loss due to the poor channel conditions. An average 15.2% of energy saving and 13% reduction in latency is achieved by the proposed scheme. Therefore, the objective for the energy-efficient low-latency communication stated in the second problem focus has been achieved.

In Chapter 5, a balanced network coded transmission scheme called BCCT has been proposed, which is developed by using random linear network coding techniques and physical layer network coding to reduce the number of required transmissions for sharing a file within the group. It uses both Wi-Fi and cellular interfaces efficiently to accomplish the data sharing quickly. The goal is to maintain the fairness among the client devices and to accomplish the process with minimum number of transmissions. The scheme utilises an information table which keeps the packet reception information and the number of transmissions each client makes in each round of data exchange process. This knowledge is important to ensure that a certain client does not run out of energy and leave the group, and for maintaining fairness among the clients. The total number of transmissions decreases while distributing the work load among the clients.

By using the physical layer network coding, the simulation results show that clients receive their required packets mostly in the earlier iteration of the process. As a result, it satisfies the requirement of participants as quickly as possible. This scheme significantly balances the number of transmissions made from each device and also reduces the total number of required transmissions by 8.92%. The subscribers gain satisfaction quickly and

reduce burden on base station.

The evaluation of the whole framework is reserved for future work of research in the practical network infrastructure. The performance of the whole framework can be predicted by the performance of each part of the framework as achieved in this dissertation. This framework is an important concept aiming to the future 5G because it depicts one of the very core parts of future wireless networks that will provide the high-rate low-energy, fast and reliable services to billions of connected devices. A practical implementation of XOR and RLNC techniques in programmable switches has been done in a doctoral work. Network coding was deployed on the switchs data plane for the first time. The evaluation was performed on Mininet network emulator. In their work, a new high-level language called P4 was used to implement the network coding functions in the programmable switches (eg. Barefoot Tofino). It is hoped that the proposed framework can also be implemented by using practical settings.

The core connection between the parts of the framework is the concept of multihop wireless communication. In every network technologies applied in this dissertation, network coding can be added as a function of transmitter or receiver or forwarder at any point along the route depending on the information such as topology, route information, and data reception information. Although the purpose can be different from one target network to another depending on the application, the overall aim of the framework is to achieve high performance operation from the perspective of multihop communication. The focus is the communication and data transfer operation in multihop wireless communication with network coding functions in a multihop scenario.

The best scenario could be a network which allows direct communication of nodes in one to two hops distance. In spite of considering the autonomous nodes in this research, the centralised control from a base station will also have positive impact for some cases such as the group data exchange. In those cases, the nodes transmit many control messages in order to reach a decision. If the central base station could take this role, there can be positive advantages for the group. This is also suitable for the cluster-based WSN where sensor nodes exchange many control messages to get a best cluster. However, due to the wide physical area, their knowledge of the whole network is limited. In this case, the help of central BS will give many advantages.

In the near future, no technology will be able to provide alone the high demand of fast and reliable communication due to the very high increasing requirements in wireless and industrial communication. The network should be flexible to integrate with other technologies. This framework is conceptualised based on the combination of future potential network technologies such as D2D, M2M and MWN to support the new requirements. The multihop wireless communication is expected to become a fundamental communication paradigm in future despite the fact that it is currently an extra option in IEEE 802.11.

The main contributions can be listed as follows:

- A new data communication framework for the predominant paradigms of future wireless communication such as 5G and IoT applications. This framework shows the incorporation of MAC protocols and network coding techniques for different data communication applications such as data transmission, data collection and data sharing by using network coding opportunity with some kinds of different topologies. This work contributes to the knowledge of how network coding can be integrated in future wireless networks to mitigate probably in some portion the ever-increasing demand of high-quality wireless communication.

- Another contribution is the design of new transmission schemes which include new MAC protocols. Transmission opportunity for the coded packets plays an important role in the decision of the performance of network coding. Therefore, a MAC protocol should reserve a priority to the encoder, in other words, the relay node as in this work, to get the transmission opportunity after the encoding operation.

- In the balanced coding and transmission scheme based on the index coding, linear network coding and physical layer network coding for the cooperative data sharing scenario, the proposed schemes consist of algorithms for the selection of next transmitter and selection of best linear combination to maintain fairness in terms of number of transmissions from each user. It is found that many control messages can be reduced by using the RLNC network coding and broadcasting. Then, the participants simply take turn to transmit their coded packets (random linear combinations) to maintain the fairness of transmissions among them.

In this dissertation, the main focus is the area of multihop wireless networks and data link layer communication protocols to achieve the performance metrics such as throughput, energy consumption, latency, number of transmissions, fairness, overhead of the protocol. The contribution of this work is only a small portion compared to the requirements of the future wireless networks, especially the current hot issues in 5G. Among the several potential techniques to contribute to the 5G network such as massive multiple input and multiple output (MIMO), cooperative communications and network coding (NC), full duplex (FD), device-to-device (D2D) communications, cognitive radio (CR), and green communications, this work falls into three categories: cooperative communications and network coding, D2D and green communication. This work has not considered the possible vulnerabilities from the security perspective yet. Security is also important for all data transmission applications. Therefore, more work is needed for the secure communication. The entire work is mainly focus on the data transmission at link level. The mechanism for interference cancellation has not yet been considered in this work for the very crowed conditions.

6.2 Future Work

Some future extension of this work could be the finding of more network coding opportunity. The current work is based on the three-node topological models, and applies only chain and triangle topologies with the MAC. The mechanism for "Y" and "X" topologies could also be considered as a future work so that the necoMAC scheme can provide more network coding service with diverse network topologies. The proposed protocol design and algorithms in the current work are based on the IEEE 802.11 wireless technologies and its specifications. In future work, it would be necessary to consider more new technologies and standards for low energy consumption.

Another possible extension could be the selection of best relay in compared to the network coder node. There would be more benefits if the relay and the network coder are identical in terms of more coding opportunity of messages. Study on the selection of a neighbour node as an helper for one's own transmission would be a useful topic for the short-range communication with plenty of connected devices in the coming near future of 5G era.

The scenario discussed in Chapter 5 could be easily extended into a data sharing group without the base station. For example, a group of friends take photos of each other and they want to share these photos to all. The task can be accomplished without the intervention of a base station. They form a local group for data sharing. One may consider privacy and security issues for this case.

One important considerable metric would be the initial battery energy percentage of a node. In BCCT, the initial energy of each client devices is not taken into account in the design of the fairness transmission scheme. The focus is on the fairness of the transmissions from the participants to accomplish the goal. However, it is possible that a client has low initial energy when it takes part in the cooperative data sharing process. In such situation, choosing such client for transmission to maintain fairness among the participants will become a burden for it. If the client is lack of energy, it will no longer be able to participate in the process and it can become a disadvantage to the other members. Therefore, considering the energy of a participant in designing the balanced scheme would be a useful extension for further research. The energy information would be useful in deciding the next transmitter and achieve more fairness to maintain the energy as long as possible.

As only binary field is used, when two combinations are added on the air by PLNC, the combination becomes a dependent combination of previously received ones. For this reason, at later iteration, the dependency is higher and BCCT/PLNC shows a small advantage. If the high-level fields are used, the problem may be solved, but the computation will also be expensive. Another method to avoid this condition is that the transmission scheme should allow both multicast transmission and PLNC transmission. By this mechanism, the addition of two combinations that will produce non-innovative packet could be reduced.

6.3 Conclusion

Based on the current situation and prospect on the future wireless communications, this research is the combination of topologies and network coding for the scenarios that have high potential for the future, communication with the help of base station and without base station. These scenarios include a mobile cloud where a direct point to point or point

to multipoint or mesh communication is allowed with the control of the base station. In this work, a WSN is utilised as an example scenario for the communication with the base station. The knowledge gained from this research is intended to support for the IoT applications. The communication from the base station and without base station is to encourage more freedom in communication with each wireless device and reduce the participation of base station to improve the scalability of the base station. This study leads to the general framework, not focusing on one specific network type. Its focus is on the topology and network coding opportunity in multihop wireless communication.

Except in multihop communication in ad hoc scenario, the communication needs the base station in WSN or cellular tower in D2D. This is because the current infrastructure has been designed in a centralised way. All communication traffic from each personal device must communicate the access point or the base station to reach the partner on the other side. This approach was invented and utilised since many years ago when the demand was not as enormous as today. Moreover, there may be problems relating to the quality of service in this current approach with the increase of connected devices in the near future.

To solve this problem, the world needs to change the way of communication from centralised to decentralised. Network coding is one of the enabling technologies in this perspective. The fundamental aspect of network coding replaces the traditional store-and-forward paradigm of communication with the compute-and-forward one. This aspect has much more advantage for the wireless communication because of the broadcast nature of wireless transmission than the wired network which can transmit even in Gbps rate. The traditional way of thinking over the overheard signal is changed to the thinking as a useful resource for encoding and decoding operation in network coding. Network coded broadcast packets can benefit many users at the same time. The implementation of network coding has already tested with the available commercial mobile phones these days.

Therefore, for the scalability issues, ever-increasing demand of quality of service and the unavoidable high overhead for control messages in the design of communication protocols, network coding is a possible solution. Based on the results and observation in this study, it could be concluded that network coding can be applied effectively by incorpo-

rating into the existing technologies, or by developing algorithms for the efficient code construction for various new scenarios to tackle some challenges of the future wireless networks such as the bandwidth and power-hungry cellular networks, Wi-Fi, and energy limited wireless sensor networks from body level to ground level, for energy and bandwidth saving and increasing throughput.

Bibliography

[1] T. Cisco, "Cisco visual networking index: global mobile data traffic forecast update the cisco ® Visual Networking Index (VNI) global mobile data traffic forecast update," The Cisco, Tech. Rep., 2017, pp. 1–29.

[2] A. Osseiran, F. Boccardi, V. Braun, K. Kusume, P. Marsch, O. Queseth, M. Schellmann, H. Schotten, H. Taoka, H. Tullberg, M. Maternia, O. Queseth, M. Schellmann, H. Schotten, H. Taoka, H. Tullberg, M. A. Uusitalo, B. Timus, and M. Fallgren, "Scenarios for the 5G mobile and wireless communications : the vision of the METIS project," *IEEE Commun. Mag.*, vol. 52, no. 5, pp. 26–35, 2014.

[3] Z. Ma, Z. Zhang, Z. Ding, P. Fan, and H. Li, "Key techniques for 5g wireless communications: network architecture, physical layer, and mac layer perspectives," *Sci. China Inf. Sci.*, vol. 58, no. 4, pp. 1–20, 2015.

[4] G. Y. Li, Z. Xu, C. Xiong, C. Yang, S. Zhang, Y. Chen, and S. Xu, "Energy-efficient wireless communications: tutorial, survey, and open issues," *IEEE Wirel. Commun.*, vol. 18, no. 6, pp. 28–35, 2011.

[5] H. Zhang, *Green communications: Theoretical fundamentals, algorithms and applications*. CRC Press, 2012.

[6] *11 internet of things (IoT) protocols you need to know about, available: Https://www.rs-online.com/designspark/eleven-internet-of-things-iot-protocols-you-need-to-know-about*, 2015.

[7] K. E. Nolan, W. Guibene, and M. Y. Kelly, "An evaluation of low power wide area network technologies for the internet of things," in *2016 Int. Wirel. Commun. Mob. Comput. Conf.*, IEEE, 2016, pp. 439–444.

[8] A. Heidarzadeh and A. Sprintson, "Cooperative data exchange with unreliable clients," *CoRR*, vol. abs/1508.03871, 2015.

[9] Y. Keshtkarjahromi, H. Seferoglu, R. Ansari, and A. A. Khokhar, "Network coding for cooperative mobile devices with multiple interfaces," *CoRR*, vol. abs/1503.02266, 2015.

[10] S. Y. E. Rouayheb, A. Sprintson, and P. Sadeghi, "On coding for cooperative data exchange," *CoRR*, vol. abs/1002.1465, 2010.

[11] Q. Zhang, J. Heide, M. V. Pedersen, and K. Rikkinen, "Network coding: Fundamentals and applications," in, M. Medard and A. Sprintson, Eds. Academic Press, 2012, ch. 5 Network Coding and User Cooperation for Streaming and Download Services in LTE Networks, pp. 115–140.

[12] G. Shen, J. Liu, D. Wang, J. Wang, and S. Jin, "Multi-hop relay for next-generation wireless access networks," *Bell Labs Tech. J.*, vol. 13, no. 4, pp. 175–193, 2009.

[13] K Loa, C c. Wu, S t. Sheu, Y Yuan, M Chion, D Huo, and L Xu, "Imt-advanced relay standards [wimax/lte update]," *IEEE Commun. Mag.*, vol. 48, no. 8, pp. 40–48, 2010.

[14] R Pabst, B. H. Walke, D. C. Schultz, P Herhold, H Yanikomeroglu, S Mukherjee, H Viswanathan, M Lott, W Zirwas, M Dohler, H Aghvami, D. D. Falconer, and G. P. Fettweis, "Relay-based deployment concepts for wireless and mobile broadband radio," *IEEE Commun. Mag.*, vol. 42, no. 9, pp. 80–89, 2004.

[15] R. Ahlswede, N. Cai, S.-Y. R. Li, and R. W. Yeung, "Network information flow," *IEEE Transactions on Information Theory*, vol. 46, no. 4, pp. 1204–1216, 2000.

[16] S. Katti, H. Rahul, W. Hu, D. Katabi, M. Médard, and J. Crowcroft, "XORs in the air: Practical wireless network coding," in *In Proc. ACM SIGCOMM*, 2006, pp. 243–254.

[17] H. Seferoglu, A. Markopoulou, and K. Ramakrishnan, "I2NC: Intra- and inter-session network coding for unicast flows in wireless networks," in *INFOCOM, 2011 Proceedings IEEE*, 2011, pp. 1035–1043.

[18] S. Chachulski, M. Jennings, S. Katti, and D. Katabi, *Trading structure for randomness in wireless opportunistic routing*, 4. ACM, 2007, vol. 37.

[19] *Network coding research group, available: Https://irtf.org/nwcrg*, 2013.

[20] P. Pahlevani, M. Hundebøll, M. V. Pedersen, D. Lucani, H. Charaf, F. P. Fitzek, H. Bagheri, M. Katz, S. M. Evice, T. O. Mart, P. Pahlevani, M. Hundebøll, M. V. Pedersen, D. Lucani, H. Charaf, H. P. Frank, H. Bagheri, and M. Katz, "Novel concepts for device-to-device communication using network coding," *IEEE Communications Magazine*, vol. 52, no. April, pp. 32–39, 2014.

[21] A. Paramanathan, M. V. Pedersen, D. E. Lucani, F. H. Fitzek, and M. Katz, "Lean and mean: Network coding for commercial devices," *IEEE Wireless Communications*, vol. 20, no. 5, pp. 54–61, 2013.

[22] S. M. Patterson, *How the open compute project's telco project could transform the iot, driverless cars*, 2016.

[23] I. . L. S. Committee *et al.*, "Wireless lan medium access control (MAC) and physical layer (PHY) specifications," *IEEE Standard*, vol. 802, no. 11, 1999.

[24] J. Gaddy, "Layer 2.5 network coding to improve 802.11 network efficiency," PhD thesis, University of Washington, 2016, pp. 1–55.

[25] M. Chaudhry and A. Sprintson, "Efficient algorithms for index coding," in *INFOCOM Workshops 2008, IEEE*, 2008, pp. 1–4.

[26] C. Fragouli, J.-Y. Le Boudec, and J. Widmer, "Network coding: an instant primer," *SIGCOMM Comput. Commun. Rev.*, vol. 36, no. 1, pp. 63–68, 2006.

[27] S. Y. R. Li, R. W. Yeung, and N. Cai, "Linear network coding," *IEEE Transactions on Information Theory*, vol. 49, no. 2, pp. 371–381, 2003.

[28] R. Koetter and M. Medard, "An algebraic approach to network coding," *IEEE/ACM Transactions on Networking*, vol. 11, no. 5, pp. 782–795, 2003.

[29] T Ho, M Medard, R Koetter, D. R. Karger, M Effros, J. Shi, and B Leong, "A random linear network coding approach to multicast," *Information Theory, IEEE Transactions on*, vol. 52, no. 10, pp. 4413–4430, 2006.

[30] A. Sprintson, "Network coding and its applications in communication networks," in *Algorithms for Next Generation Networks*, Springer, 2010, pp. 343–372.

[31] P. A. Chou and Y. Wu, "Network coding for the internet and wireless networks," *IEEE Signal Processing Magazine*, vol. 24, no. 5, pp. 77–85, 2007.

[32] J. Qureshi, C. H. Foh, and J. Cai, "Online XOR packet coding: efficient single-hop wireless multicasting with low decoding delay," *Computer Communications*, vol. 39, pp. 65–77, 2014.

[33] M. Z. Farooqi, S. M. Tabassum, M. H. Rehmani, and Y. Saleem, "A survey on network coding: from traditional wireless networks to emerging cognitive radio networks," *Journal of Network and Computer Applications*, vol. 46, pp. 166–181, 2014.

[34] J Le, J. C. S. Lui, and D.-M. Chiu, "DCAR: distributed coding-aware routing in wireless networks," *IEEE Trans. Mob. Comput.*, vol. 9, no. 4, pp. 596–608, 2010.

[35] M. A. Iqbal, B. Dai, B. Huang, A. Hassan, and S. Yu, "Survey of network coding-aware routing protocols in wireless networks," *J. Netw. Comput. Appl.*, vol. 34, no. 6, pp. 1956–1970, 2011.

[36] Z Jian, Y. P. Chen, and I Marsic, "BEND: MAC-layer proactive mixing protocol for network coding in multi-hop wireless networks," in *Proceedings of the 9th ACM International Symposium on Mobile Ad Hoc Networking and Computing*, vol. 1, 2008, pp. 447–448.

[37] X. Wang, S. S. Member, J. Li, S. S. Member, and F. Tang, "Network coding aware cooperative MAC protocol for wireless ad hoc networks," *IEEE Transactions on Parallel and Distributed Systems*, vol. 25, no. 1, pp. 167–179, 2014.

[38] S. C. Liew, S. Zhang, and L. Lu, "Physical-layer network coding: Tutorial, survey, and beyond," *CoRR*, vol. abs/1105.4261, 2011.

[39] P. Ostovari, J. Wu, and A. Khreishah, "Network coding techniques for wireless and sensor networks," in *The Art of Wireless Sensor Networks: Volume 1: Fundamentals*, H. M. Ammari, Ed. Berlin, Heidelberg: Springer Berlin Heidelberg, 2014, pp. 129–162.

[40] N. R. Anjana, S. Valiveti, S. Garg, and K. Kotecha, "Topology management in ad hoc network," *International Journal of Computer Applications*, vol. 10, no. 8, pp. 1–5, 2010.

[41] P. Lu, X. Wang, Y. Tan, and A. O. Lim, "Network coding-aware MAC for throughput optimization in wireless multihop networks," *IPSJ SIG Mobile Computing and Ubiquitous Communications (SIG-MBL 63), Tokyo, Japan*, vol. 2012-MBL-63, no. 14, 2012.

[42] S. Abdel Hamid, H. S. Hassanein, and G. Takahara, "Routing for wireless multi-hop networks," in *Routing for Wireless Multi-Hop Networks*, ser. SpringerBriefs in Computer Science, New York, NY: Springer New York, 2013, pp. 11–24.

[43] H. R. B. Awerbuch D. Holmer, "High throughput route selection in multi-rate ad hoc wireless networks," *Lecture Notes in Computer Science*, 2003.

[44] S. Kumar, V. S. Raghavan, and J. Deng, "Medium access control protocols for ad hoc wireless networks: a survey," *Ad Hoc Networks*, vol. 4, no. 3, pp. 326–358, 2006.

[45] J. F. Kurose and K. W. Ross, *Computer Networking A Top-Down Approach*, 5. 2013. arXiv: 8177588788.

[46] A. O. Lim and S. Yoshida, "A 2-hop path selection protocol (2PSP) in multi-rate ad hoc wireless networks," *IEICE Trans. Commun.*, vol. 2007-E90-B, no. 1, 2007.

[47] L. M. Feeney and M. Nilsson, "Investigating the energy consumption of a wireless interface in an ad hoc networking environment," *INFOCOM 2001.Twentieth Annual Joint Conference of the IEEE Computer and Communication Societies Prodeedings. IEEE*, vol.3, pp.1548 - 1557, April 2001.

[48] T. Braun, A. Kassler, M. Kihl, V. Rakocevic, V. Siris and G. Heijenk, *Multi-Hop Wireless Networks*. Springer Verlag, USA, 2009, vol. 31, pp. 1–68.

[49] R. Palacios-Trujillo, B. H. Dabi, J. Alonso-Zarate, F. Granelli, F. H. P. Fitzek, and N. L. S. da Fonseca, "Network coding-aware IEEE 802.11 MAC protocol using batch transmissions and multiple reverse direction exchanges," in *2016 IEEE International Conference on Communications (ICC)*, IEEE, 2016, pp. 1–7.

[50] "ISO/IEC Standard for Information Technology - Telecommunications and Information Exchange Between Systems - Local and Metropolitan Area Networks - Specific Requirements Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications (Includes IEEE Std 802.11, 1999 Edition; IEEE Std 802.11A.-1999; IEEE Std 802.11B.-1999; IEEE Std 802.11B.-1999/Cor 1-2001;

and IEEE Std 802.11D.-2001)EC standard for information technology - telecommunications and information exchange between systems - local and metropolitan area networks - specific requirements part 11: Wireless LAN medium access control (MAC) and physical layer (PHY) specifications (includes IEEE std 802.11, 1999 edition; IEEE std 802.11a.-1999; IEEE std 802.11b.-1999; IEEE std 802.11b.-1999/cor 1-2001; and IEEE std 802.11d.-2001)," *ISO/IEC 8802-11 IEEE Std 802.11 Second edition 2005-08-01 ISO/IEC 8802 11:2005(E) IEEE Std 802.11i-2003 Edition*, pp. 1–721, 2005.

[51] H. C. Lee, "A MAC throughput in the wireless LAN," *InTech*, pp. 1–61, 2012.

[52] M Hay, B Saeed, C Lung, T Kunz, and A Srinivasan, "Network coding and quality of service for mobile ad hoc networks," *International Journal of Communications, Network and System Sciences*, vol. 7, pp. 409–422, 2014.

[53] H. Zimmermann, "OSI reference model–the ISO model of architecture for open systems interconnection," *IEEE Trans. Commun.*, vol. 28, no. 4, pp. 425–432, 1980.

[54] C. H. Huang, D. Umehara, S. Denno, M. Morikura, and T. Sugiyama, "Enhancement of CSMA/CA and network coding in single-relay multi-user wireless networks," *IEICE Transactions on Communications*, vol. E93-B, no. 12, pp. 3371–3380, 2010.

[55] P. Pahlevani, D. E. Lucani, M. V. Pedersen, and F. H. P. Fitzek, "PlayNCool: opportunistic network coding for local optimization of routing in wireless mesh networks," in *2013 IEEE Globecom Workshops, GC Wkshps 2013*, IEEE, 2013, pp. 812–817.

[56] W. P. Wang, W. Wu, Q. J. Guan, and J. X. Wang, "TCAR: a new network coding-aware routing mechanism based on local topology detection," *Journal of Central South University*, vol. 21, no. 8, pp. 3178–3185, 2014.

[57] G.-j. Arnaud, Y. Zhao, Z. Dong, M. Iwai, K. Sezaki, and Y. Tobe, "An extended network coding opportunity discovery scheme in wireless networks," *International Journal of Computer Networks & Communications (IJCNC) Vol.4, No.1, January 2012*, vol. 4, no. 1, pp. 63–77, 2012.

[58] R. Palacios, F. Granelli, A. Paramanathan, J. Heide, and F. H. P. Fitzek, "Coding-aware MAC: providing channel access priority for network coding with reverse direction DCF in IEEE 802.11-based wireless networks," *2014 IEEE Int. Conf. Commun. ICC 2014*, pp. 1272–1277, 2014.

[59] Jangeun Jun, P. Peddabachagari, and M. Sichitiu, "Theoretical maximum throughput of ieee 802.11 and its applications," in *Second IEEE Int. Symp. Netw. Comput. Appl. 2003. NCA 2003.*, IEEE Comput. Soc, 2003, pp. 249–256.

[60] Yang Xiao and J. Rosdahl, "Throughput and delay limits of ieee 802.11," *IEEE Commun. Lett.*, vol. 6, no. 8, pp. 355–357, 2002.

[61] A. Iyer, C. Rosenberg, and A. Karnik, "What is the right model for wireless channel interference?" *IEEE Trans. Wirel. Commun.*, vol. 8, no. 5, pp. 2662–2671, 2009.

[62] T. W. Kuo and M. J. Tsai, "On the construction of data aggregation tree with minimum energy cost in wireless sensor networks: NP-completeness and approximation algorithms," in *Proceedings - IEEE INFOCOM*, 2012, pp. 2591–2595.

[63] J. Gao, L. Guibas, N. Milosavljevic, and J. Hershberger, "Sparse data aggregation in sensor networks," *Information Processing In Sensor Networks*, vol. Proceeding, pp. 430–439, 2007.

[64] C. Luo, F. Wu, J. Sun, and C. W. Chen, "Compressive data gathering for large-scale wireless sensor networks," *Proceedings of the 15th annual international conference on Mobile computing and networking (MobiCom '09)*, no. 800, p. 145, 2009.

[65] E. Fasolo, M. Rossi, J. Widmer, and M. Zorzi, "In-network aggregation techniques for wireless sensor networks: a survey," *IEEE Wireless Communications*, vol. 14, no. 2, pp. 70–87, 2007.

[66] S Lindsey, C Raghavendra, and K. M. Sivalingam, "Data gathering algorithms in sensor networks using energy metrics [pegasis]," *IEEE Transactions on Parallel and Distributed Systems*, vol. 13, no. 9, pp. 924–935, 2002.

[67] G. Gupta and M. Younis, "Load-balanced clustering of wireless sensor networks," *IEEE International Conference on Communications, 2003. ICC '03.*, vol. 3, pp. 1848–1852, 2003.

[68] O. Younis, "Distributed clustering in ad-hoc sensor networks: a hybrid, energy-efficient approach [heed]," *Third AnnualJoint Conference of the IEEE*, pp. 1–12, 2004.

[69] P. Salija and B. Yamuna, "Optimum energy efficient error control techniques in wireless systems: A survey," *Journal of Communications Technology and Electronics*, vol. 60, no. 11, pp. 1257–1263, 2015.

[70] W. B. Heinzelman, A. P. Chandrakasan, and H. Balakrishnan, "An application-specific protocol architecture for wireless microsensor networks," *IEEE Transactions on Wireless Communications*, vol. 1, no. 4, pp. 660–670, 2002.

[71] S. Mohammad, "Opportunistic cooperative communication using buffer-aided relays," *Qatar Foundation Annual Research Forum Proceedings*, ICTP 025, 2013.

[72] ScienceDaily, "Cooperative driving will become common: Data exchange between vehicles and road network," Technical Research Centre of Finland (VTT), Science News from research organizations, 2015.

[73] T. Courtade and R. Wesel, "Coded cooperative data exchange in multihop networks," *Information Theory, IEEE Transactions on*, vol. 60, no. 2, pp. 1136–1158, 2014.

[74] Y. Sui, X. Wang, J. Wang, L. Wang, and S. Hou, "Deadline-aware cooperative data exchange with network coding," *Computer Networks*, vol. 97, pp. 88 –97, 2016.

[75] S. Tajbakhsh and P. Sadeghi, "Coded cooperative data exchange for multiple unicasts," in *Information Theory Workshop (ITW), 2012 IEEE*, 2012, pp. 587–591.

[76] Q. Zhang, F. Fitzek, and M. Katz, "Cooperative power saving strategies for IP-services supported over DVB-H networks," in *Wireless Communications and Networking Conference, WCNC 2007. IEEE*, 2007, pp. 4107–4111.

Publications

Journals

[1] Nyan Lin, Tan, Y. and Lim Y., "Efficient Network Coding Based Data Transfer Framework for Multihop Wireless Communication," *Int. J. of Networking and Virtual Organisations*, 2018. (To be published)

[2] Nyan Lin, B. M. Kurkoski, Y. TAN and Y. LIM, "Data Sharing among Wireless Client Devices in Cooperative Manner with Minimum Transmissions," *Journal of Ambient Intelligence and Humanized Computing*, 2018. (To be published)

[3] Nyan Lin, Y. Tan, and Y. Lim, "Network Coding-aware Medium Access Control (necoMAC) Scheme for Multihop Wireless Networks," *Journal of Wireless Communications*, 2017. (Under Review)

International Conferences

[4] Nyan Lin, R. Ngeth, Y. Tan, and Y. Lim, "Energy Efficient Data Gathering with Network Coding for Cluster-based Wireless Sensor Networks," in *Proceedings of the 14th International Conference on Computer Applications (ICCA)*, pp. 8-13, Yangon, Myanmar, February 2016.

[5] Nyan Lin, R. Ngeth, K. Sriviriyakul, Y. Tan and Y. Lim, "necoMAC: Network Coding aware MAC Protocol for Multirate Wireless Networks," in *2016 IEEE 30th International Conference on Advanced Information Networking and Applications (AINA)*, pp. 384-391, Switzerland, March 2016.

[6] Nyan Lin, B. M. Kurkoski, Y. Tan, and Y. Lim, "Balanced Cooperative Linear Data Exchange with Physical Layer Network Coding," in *Proceedings of the International Conference on Internet of Things and Cloud Computing (ICC)*, pp 1-6, Cambridge, UK, March 2016.

Local Conferences

[7] Nyan Lin, Y. Tan, and Y. Lim, "Study of Energy Consumption for Data Gathering in NC-based Wireless Sensor Networks," in *Proceedings of the IEICE General Conference*, vol. 2, pp. 155, Ritsumeikan University, Kyoto, February 2015.

[8] Nyan Lin, K. Sriviriyakul, R. Ngeth, Y. Tan, and Y. Lim, "Study of Network Coding-aware MAC protocol for Multihop Wireless Networks," in *Proceedings of the Society Conference of IEICE*, vol. 2, pp. 67, Tohoku University, Sendai, August 2015.

[9] Nyan Lin, R. Ngeth, Y. Tan, and Y. Lim, "Study of Balanced Cooperative Coding Scheme for Linear Data Exchange Problem," in *Proceedings of the IEICE General Conference*, vol. 2, pp. 176, Kyushu University, Fukuoka, March 2016.

[10] Nyan Lin, Yasuo TAN and Yuto LIM, "Study of Network Coding-based MAC Schemes for Different Topologies in Wireless Networks," in *Proceedings of the IEICE ASN Workshop*, vol. ASN2017-99, pp. 91-96, Oita, January 2018.

[11] R. Ngeth, Nyan Lin, Y. Tan and Y. Lim, "Network Coding Scheme based on Selection of Number Input Packets in X-topology Network," in *Proceedings of the IEICE General Conference*, vol. 2, pp. 177, Kyushu University, Fukuoka, March 2016.

Appendix A

Upper and lower bounds for BCCT

In this section, the upper and lower bounds on the number of transmissions from [10] are described.

The minimum number of transmissions is greater or equal to $n - n_{min}$ where n is the number of packets and n_{min} is the minimum number of packets held by a client, i.e., $n_{min} = \min_{1 \leq i \leq k} n_i$. If all clients initially have the same number of packets $n_{min} < n$, i.e., $n_i = n_{min}$ for $i = 1, \cdots, k$ clients, then the minimum number of transmissions is greater than or equal to

$$n - n_{min} + 1 \tag{A-1}$$

The upper bound on the minimum required number of transmissions is less than or equal to

$$\min_{1 \leq i \leq k} \{|\overline{X_i}| + \max_{1 \leq j \leq k} |\overline{X_j} \cap X_i|\} \tag{A-2}$$

for $|\mathbb{F}| \geq k$. Each client c_i initially holds a subset X_i of packets in $X = \{x_1, \cdots, x_n\}$, i.e., $X_i \subseteq X$. And $n_i = |X_i|$ denotes the number of packets initially available to client c_i, and $\overline{X_i} = X \setminus X_i$.