

## CHAPTER ONE

### INTRODUCTION

Nowadays, the Internet is mostly everywhere as it has reached many parts of the world, even most of the remotest places on the planet. People are connected to their smart devices most the time. If you think you cannot live without your Smartphone, make sure you are not alone. The extensive development of new communication technologies and the needs of these technologies is the motivation for a large spectrum of networking researchers who seek to enhance communication systems hardware and software infrastructure. Rapidly over the past two decades, spacious communications applications and services have been developed, especially in wireless communications.

In the past decade, wireless networks have invaded the whole world with their extensive applications and services. Nowadays, the number of cellular phones and their products exceeds wired phones and the difference is growing. Wireless technologies have had a huge global impact in a short time, impacting on human life in many ways including healthcare services, environmental protection, business communications and services, news, entertainment, education, and social life, among others. New cellular technologies have been introduced, such as the early 0G to the latest technology the 4G services with native-IP networks. Unfortunately, most of these new technologies have not shown their expected potential.

Most of the current Internet traffic is managed and controlled by the Transmission Control Protocol (TCP). The TCP traffic uses a window-based congestion control mechanism. This mechanism tries to preserve each TCP flow's sending rate equal to the flow's fair share of the links capacity, and also responds to link congestion by reducing the flow sending rate. The TCP congestion control mechanisms are well designed and optimized to perform well over wire links. However, these mechanisms experience severe performance degradation over wireless links. The poor performance of TCP over wireless links has motivated many networking researchers to improve and enhance its mechanisms.

This chapter is organized as follows. The first section gives a brief introduction to the wireless networks generations. In section two, we introduce the concept of network layers and the ISO-OSI model. An overview of the TCP over wireless links will be found in section three. Section four introduces the research problem statement. The research objectives will be listed in section five. The research scope and limitations will be listed in section six. Section seven lists the significant of the study. In section eight the research methodology will be introduced. Finally, section nine provides the outline of the thesis.

## 1.1 Wireless technology and LTE

In 1947, the concept of a cell was introduced by Bell Labs, but the new technology attracted few users at that time, due to the high cost of implementing such networks (Sesia & Toufik, 2011). In 1980, the first mobile generation was introduced in large-scale commercial communication system. Since that time, several mobile generations have been developed up to the latest technology the 4G LTE and LTE-Advance networks.

The development of cellular technologies can be traced in sequence, but different generations from the early the 0G services like Mobile Telephone Service (MTS) and Improved Mobile Telephone Service (IMTS), to the 1G analogue cellular networks, to the 2G digital cellular networks, to the WiMAX and the broadband data service in the 3G, and the latest technology the 4G services with native-IP networks, including LTE and LTE-Advance.

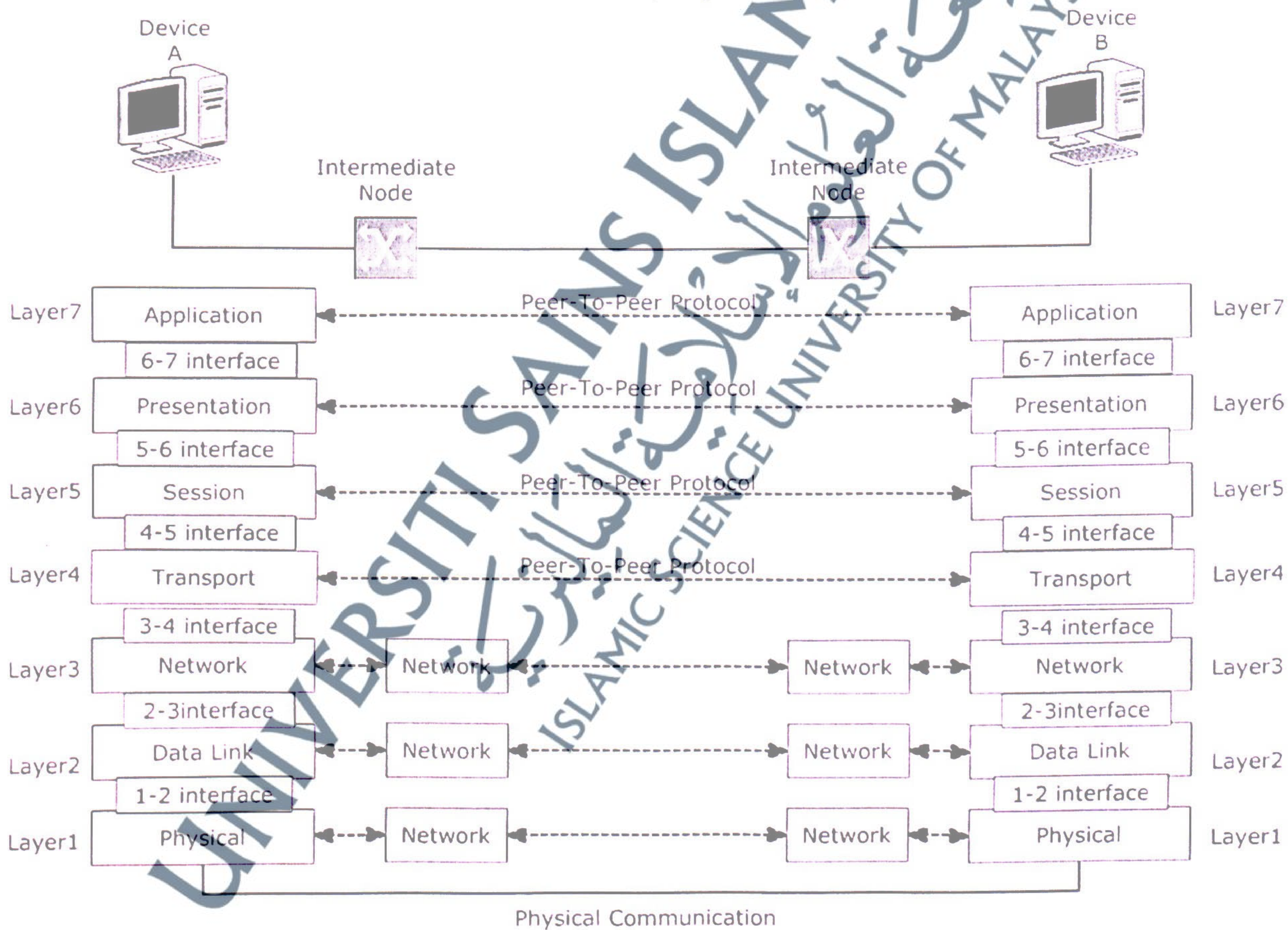
The first version of LTE was introduced with the release of 8.0 of the third Partnership Project (3GPP) specifications series in 2008. LTE is the leading of IP-based Orthogonal Frequency-Division Multiple Access (OFDMA) wireless mobile technology, in response to the need to adapt to increasing data usage by new applications. LTE uses radio access called Evolved UMTS Terrestrial Radio Access Network (E-UTRAN). 3GPP specifies the Evolved Packet Core (EPC) network architecture to support the E-UTRAN through a reduction in the number of network elements, simpler functionality, improved redundancy, and ability to allow connections and hand-over to another fixed line, with wireless access technologies to deliver a seamless mobile experience.

## 1.2 Network layers

Computer networks are developed and designed using different entities. These heterogeneous networks can communicate with one another using the networks' standards. The ISO-OSI model and the Internet model are the most widely known standards. The OSI model was proposed as the standard for all the communication networks from 1970 until 1990. Nowadays, the TCP/IP or Internet model is the standard framework used to design communication network systems. Chapter 2 provides comprehensive details about the TCP/IP model.

The International Standards Organization (ISO) was established in 1947. It is a multinational organization dedicated to worldwide agreement on standards. The ISO introduced the network communication Open System Interconnection (OSI) model in 1970 (Forouza & Fegan, 2007). The OSI model is not a protocol; it is a framework for understanding and designing heterogeneous networks to be flexible, robust, and interoperable. The OSI model is layered into seven separate but related layers; each layer is a part of the process of sending and receiving data between networks nodes. Figure 1.1 shows the seven layers of the OSI model ordered from top to down.

**Figure 1.1:** OSI Model Layers Interaction (Adopted from Forouza & Fegan, 2007)



- **Application Layer:** this layer is responsible for providing the network services to the end-user's applications using many application protocols

including, HyperText Transfer Protocol (HTTP), File Transfer Protocol (FTP), and Simple Network Management Protocol (SNMP), among others.

- **Presentation Layer:** this layer is responsible for the syntax and semantic information exchange. Three main functionalities are provided by this layer:
  1. Translation: uses a common format of encoding between two different computers encoding systems.
  2. Encryption: transform the original data to another format to ensure privacy.
  3. Compression: to reduce the size of the sent data.
- **Session Layer:** provides the services of a dialog controller, and synchronization. Dialog controlling includes determining the mode of communication between the connection peer by either half-duplex (two way but not at the same time) or full-duplex (two way at the same time).
- **Transport Layer:** this layer is responsible for process-to-process intact and in-order delivery of packets. The transport layer is concerned with other functionalities, including:
  1. Service-point addressing: each service has its unique address in which it can communicate with other services in different computers. The transport layer header must handle such addresses which are called port address to ensure the process to process delivery.
  2. Segmentation and reassembly: this divides the messages into small segments and attaches a sequence number with each segment to ensure reliability.
  3. Connection control: transport protocol can be either connection-oriented or connectionless. Connection-oriented transport protocols establish a connection between the communicated peers before sending any segment and terminating this connection after sending all the data. On the other hand, connectionless protocols send each segment as an independent message to its destination.

4. Flow control: controls the sending rate to ensure that it will not overwhelm the destination buffers.
  5. Error control: ensures that the entire packet is delivered correctly, with no damage, loss, or duplication.
  6. Congestion control: limits the sending rate to the links capacity.
- **Network layer:** this is responsible for delivery packets from the source-to-destination across two different networks. Two main functions are performed in this layer:
    1. Logical addressing: attaches the logical addresses of both the source and the destination to the header of each packet.
    2. Routing: provides the mechanism of routing and switching the packets to their final destination across the network.
  - **Data Link Layer:** this layer is responsible for transferring frames from hop-to-hop across the network to the destination. This process involves:
    1. Framing: dividing the bit streams received from the network layer to frames.
    2. Physical addressing: attaching the physical addresses of both the source and the next hop to the header of the frames.
    3. Flow Control: controlling the sending rate to match the receiver's processing rate.
    4. Error Control: attaching a trailer to the frames to recognize loss, damage, and duplicate frames.
  - **Physical Layer:** this layer deals with the mechanical and electrical specifications of the transmission medium to coordinate the functions required to carry a bit stream over this medium. This process involves:
    1. Defining the physical specifications of interfaces and mediums.
    2. Representing bit streams: how to change the streams bits into signals (electrical or optical).
    3. Defining the transmission data rate.
    4. Defining the physical topology.

5. Determining the transmission mode: simplex, half-duplex, or full-duplex.

### 1.3 TCP/IP

The first outline for the Transfer Control Protocol (TCP) was given in 1974, by Bob Kahn and Vinton Cerf. They continued working on it during the 1970s, while the ARPAnet, a network developed in the United States intended to link research facilities, was starting to grow. This ARPAnet first used the Network Control Protocol (NCP) to operate the network, but in 1983 it was replaced by the TCP in combination with the Internet Protocol (IP). The TCP/IP took care of the routing of messages in the network. This ARPAnet started to grow and also became available to home users, and soon the public part of this network was called the 'Internet'. The popularity of the Internet also contributed to the popularity of the TCP/IP, which is one of the most widely used network protocols today (Comer, 2015, p.430).

The TCP achieves a seemingly impossible task, as it uses the unreliable Internet Protocol (IP) datagram service to provide a reliable data delivery service. The TCP provides a reliable, connection-oriented, full-duplex stream transport service which allows two application programs (connection peer) to exchange data in either direction and then terminate the connection: "Each TCP connection is started reliably and terminated gracefully" (Comer, 2015, p.431).

The TCP/IP is the protocol stack used on the internet, usually called the Internet Protocol suite. The TCP/IP suite is a five layer model: from the top application, transport, network, data link, and at the bottom the physical layer. Each

layer has a series of protocols to provide services needed from the top layer and this can be done through the service interfaces between them.

The emergence of the rapid growth of computer network applications and services requires reliable network environments. Excessive traffic between network nodes reduces the throughput due to overwhelming the network resources. Therefore, the TCP improves many mechanisms to ensure its reliability. Firstly, the TCP maintains a sequencing mechanism to handle duplicates and out-of-order packets. Secondly, it provides a retransmission mechanism to handle lost and damaged packets. Thirdly, the TCP has developed techniques to avoid replay in case of long delay. Fourthly, it provides flow control techniques to limit the sending rate to the receiver's processing rate, and finally it provides congestion control mechanisms to limit the TCP sending rate to match the network resource capacity.

#### 1.4 Problem statement

Wireless technologies have experienced intensive improvements in the past two decades. Expensive applications and services have been developed. Therefore, it has become an urgent need to make the traditional techniques adapt to the new technologies to show their expected potential. Spacious new technologies have been introduced, especially in wireless communication. The latest was the Long Term Evolution (LTE), marketed as 4G. Unfortunately, LTE has not shown its expected potential due to the use of classical techniques which do not support its new requirements.

The TCP has not properly supported the mobility features of the advanced wireless networks, especially the LTE networks (Amin & Yla-Jaaski, 2013; Ahn et al.,

2012). Moreover, TCP performance is seriously degraded when the core network suffers congestion because of heavy retransmission due to timeout expirations (Abed et al., 2013; Abed et al., 2013; Yahia, 2012). The TCP assumes that any packet loss is an indication of a predicted congestion. However, wireless networks are characterized by: a high bit-error rate, sporadic connectivity, long delays, and handovers. The TCP cannot recognize the difference between real congestion losses and other types of loss (Maisuria & Patel, 2012; Park & Jang, 2011).

In addition, to achieve a seamless mobility across radio cells, LTE performs a hard handover algorithm to deliver a mobile user from one base station to another. The faster the mobile user is moving, the more handover occurs, making the TCP handle more traffic and potentially causing serious congestion or even network failure (Abed et al., 2012; Lou et al., 2015; Wang et al., 2012).

Several studies for improving TCP performance over wireless networks have been conducted (Park & Jang, 2011). However, even though most of the proposed methods and techniques have shown noticeable improvements, the drawback is that most of them are prediction-based techniques which require very complex functions at the mobile station. Furthermore, some methods (Racz et al., 2007; Pacifico et al., 2009; Liu & Lee Y. B. , 2011) require changes to the standard LTE handover procedure to support them (Amin & Yla-Jaaski, 2013; Ahn et al., 2012).

In some of the above, LTE has not yet shown its expected potential performance. Many assumptions of the wired networks are not applicable to the wireless networks because of the loss nature of the wireless networks environments, in which most of the losses are due to network failure and not to real congestion (Abed et al., 2013; Yahia , 2012; Maisuria & Patel, 2012). Therefore, the TCP needs to adapt to

the wireless environment to recognize real congestion from other types of loss (Maisuria & Patel, 2012; Luo et al., 2015). Other improvements could be added to the TCP to extend the fairness issue over the network's resources (Yahia , 2012).

## 1.5 Research objectives

The aim of this study is to evaluate a new congestion control mechanism to enhance the TCP over common wireless networks, and the LTE networks in particular. The proposed congestion control mechanism will enhance the TCP as a transport protocol to recognize the congestion losses from other sources of loss. This aim will lead to the following main objectives:

1. To establish baseline knowledge of TCP performance over wireless links, firstly, by addressing the TCP congestion control mechanisms issues over wireless links, and secondly by analysing the widely used TCP congestion control algorithms and presenting the latest enhanced algorithms.
2. To design an enhanced End-to-End TCP congestion control mechanism to be integrated with the TCP in the TCP/IP transport layer.
3. To model and evaluate the new enhanced congestion control mechanism over wireless links and especially, over LTE network. The new enhanced mechanism is compared to the common End-to-End TCP variants using a simulation-based evaluation method.

In addition, there are other control objectives to the proposed congestion control mechanism, listed as follows:

1. Avoidance of network overload: to reduce the probability of packets losses due to congestions.

2. Efficient resource utilization: keep all the links bottleneck full utilized.
3. Preservation of fair sharing: one of the main congestion control algorithms evaluating metrics is fairness. Therefore, the new modified algorithm should preserve a fair resource sharing between different TCP flows.
4. Preservation of the End-to-End TCP principle.
5. Avoidance of packet retransmission unless a clear congestion loss is recognized.
6. Enhancement of the TCP over large delay, high capacity networks. Current TCP congestion control mechanisms take large Round Trip Time ( $RTT$ ) for the window size to grow as the bandwidth-delay product. If the  $RTT$  is already large, then the congestion control algorithm leads to poor performance.

## 1.6 Research scope and limitations

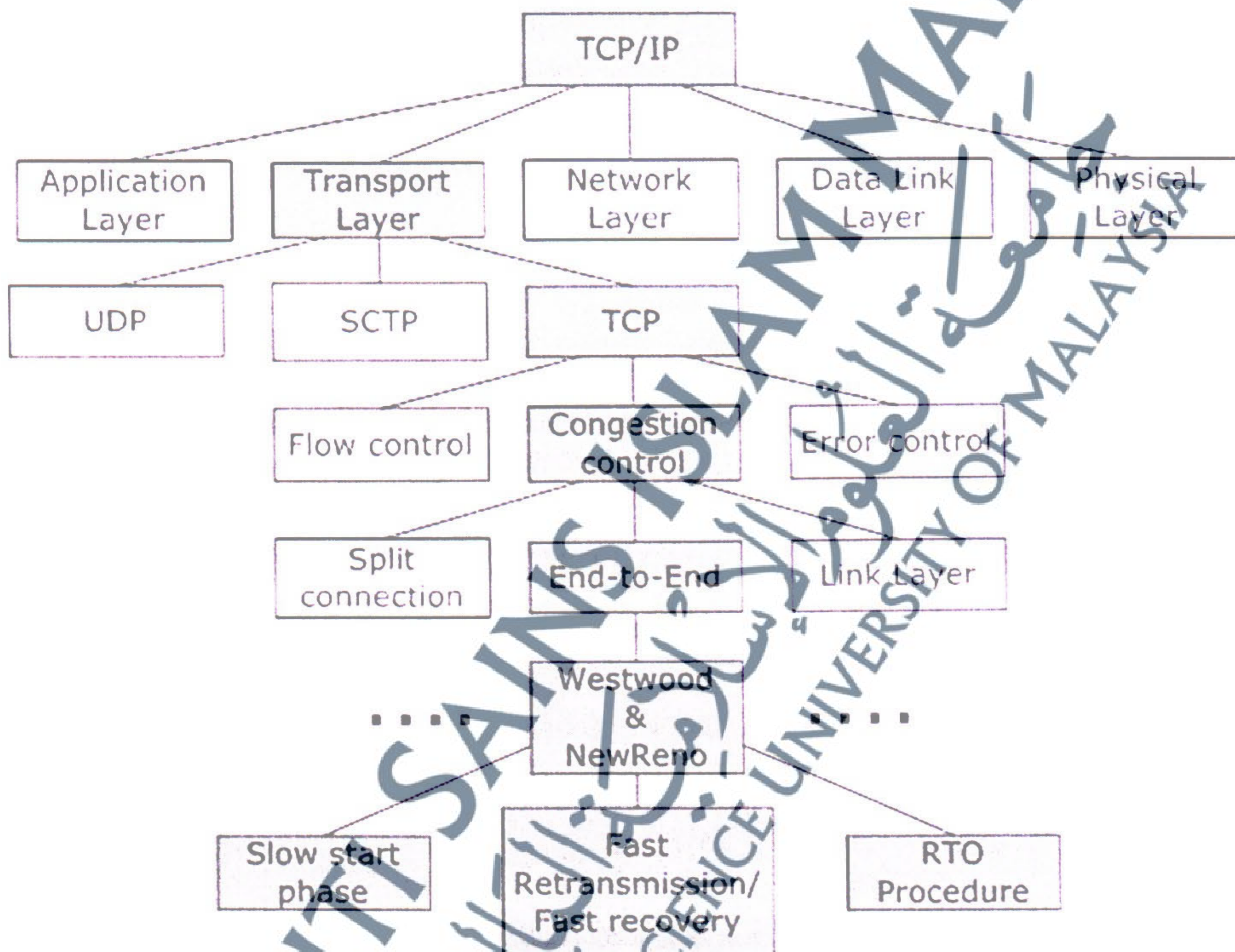
This research is focused on enhancing TCP congestion control mechanisms over wireless links, mainly, over LTE networks. Particularly, this research concentrates on modifying the TCP NewReno slow start phase, fast retransmission fast recovery algorithm, and retransmission procedure based on the TCP Westwood bandwidth estimation method. The highlighted boxes in figure 1.2 indicate the areas of the research.

On the other hand, the limitations and the assumptions of the research are:

- The proposed solution will be compared with limited TCP variants as the ns-3 simulation supports.
- Traffic is assumed to be File Transfer Protocol (FTP) traffic.

- Packet Error Rate is assumed to be uniformly distributed as the simulation supports.
- The receiver window is large enough to make sure it will not affect the sender sending rate.
- The LENA model is used to represent the LTE simulation model.

Figure 1.2: The research scope



## 1.7 Significance of the study

The significance of the study includes:

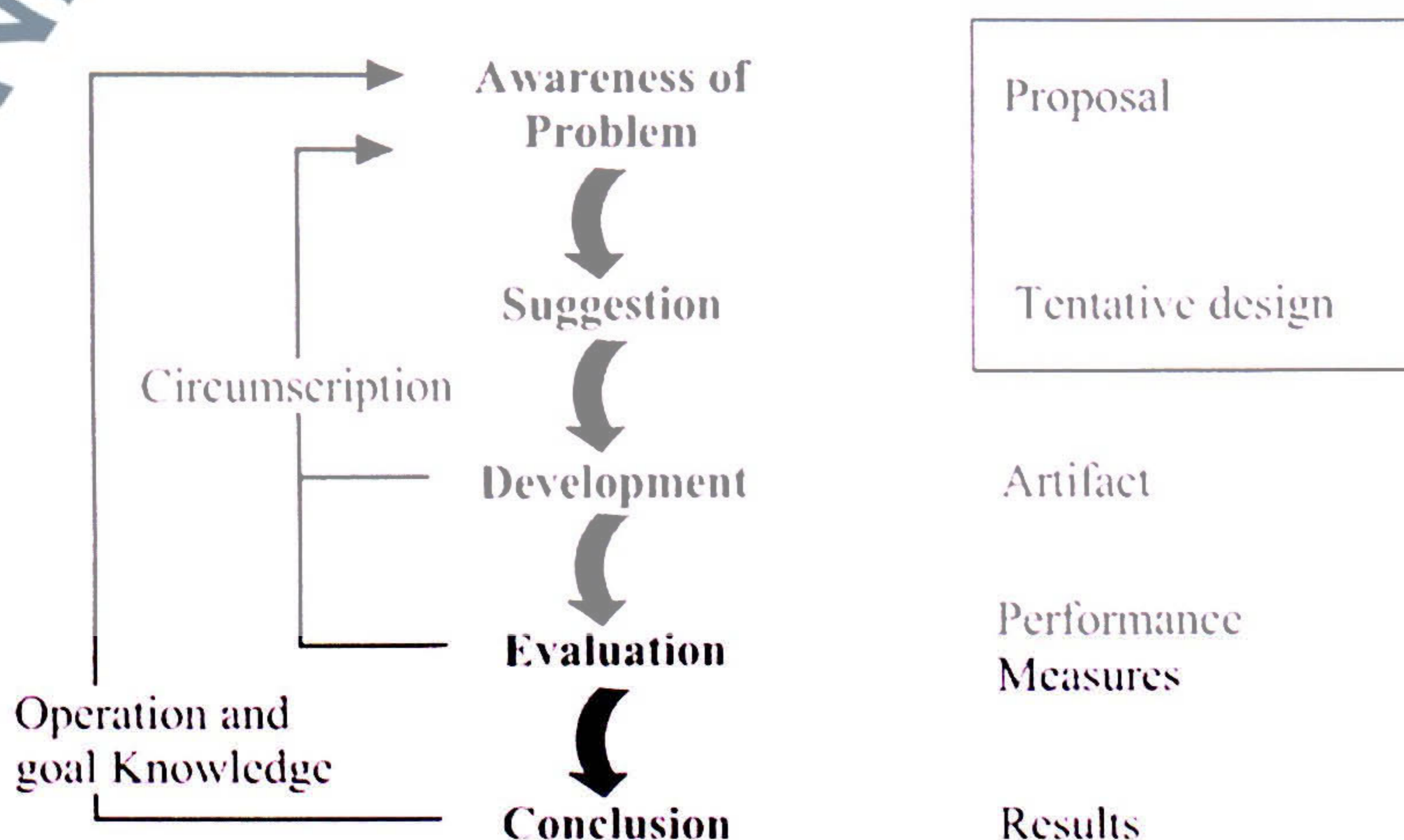
1. A baseline performance evaluation model of TCP Tahoe, TCP Reno, TCP NewReno, and TCP Westwood as a transport protocols over LTE networks.
2. The design of an enhanced End-to-End TCP congestion control mechanism based on NewReno and Westwood TCP variants. This enhanced includes:

- A new method to properly set the initial Slow Start threshold value based on the current bandwidth capacity.
  - A new method to accelerate the Slow Start exponential increment based on the estimated bandwidth.
  - A new method to enhance the Fast Retransmission mechanism.
  - A new method to enhance the Fast Recovery mechanism.
  - A new method to enhance TCP Retransmission Timeout procedure.
3. The implementation and analysis of the enhanced congestion control algorithms.
  4. A comparison of the proposed congestion control algorithm to TCP NewReno, and TCP Westwood.

## 1.8 Research Methodology

After investigating numerous research methodologies, the researcher decided to adopt the General Methodology of Design Research because it fulfils the requirements of this research. The general methodology of design research contains five main phases as shown in figure 1.3.

Figure 1.3: The general methodology of design science research



**Awareness of the problem phase:** The main objectives of this phase are to observe and identify the area of the research problem. In this phase, the researcher defines the research problem by investigating the current work in the area of TCP congestion control algorithms over wireless networks, and this specifying the objectives and scope of the study.

**Suggestion phase:** This phase includes the analysis of the problem and decomposes the complex problem into small and understandable components. The main purpose of this phase is gathering the literature review in the research area, reviewing related works, and defining the research methodology.

**Development phase:** In this phase, the proposed algorithms will be modelled in the format of Finite State Machine (FSM) to describe the behaviour of the modified TCP congestion control algorithms. Moreover, this phase includes constructing the proposed algorithms using C++ programming language. It also includes debugging and verification of artifact functionality using black box techniques.

**Evaluation phase:** Once the artifact is constructed and its functionality is verified, the evaluation phase will be started. For the purposes of this study, the simulation methods will be used to evaluate the usability of the proposed algorithm and ensure that the objectives of the study have been accomplished. Network Simulator ns-3 will be used to test and evaluate the proposed algorithms. The results will be compared with the current TCP variations to ensure the improvements of the proposed algorithms.

## 1.9 Thesis outlines

This thesis is organized in six chapters. The remainder of this thesis is organized as follows:

Chapter 2 presents a background study on common transport protocols in communication systems. Firstly, it presents a brief introduction on layering concepts, and then TCP fundamentals are discussed. After this, we present a detailed study on the congestion control mechanisms of common TCP variants. Next we present the literature on TCP congestion control mechanisms over wireless networks, and finally we introduce the SCTP protocol.

Chapter 3 presents the research methodology used in this study. Firstly, it introduces the performance evaluation techniques in computer networks. Then, it details the selected network simulator. The rest of the chapter discusses the simulation experimental design, experimental hardware and software, validation and verification, and the network performance evaluation metrics.

Chapter 4 introduces a new End-To-End congestion control mechanism based on TCP NewReno and TCP Westwood. Three new modifications to TCP NewReno are presented in this chapter. Moreover, this chapter presents the evaluation of the proposed new modifications by comparing their performance with TCP Westwood and TCP NewReno performance in certain network performance metrics.

Chapter 5 discusses implementing and evaluating the proposed modifications in chapter four over LTE networks. Moreover, this chapter will introduce the LTE technology, implementing LTE in Network Simulator-3 (ns-3), and TCP performance over LTE. Then, we discuss the implementation of TCP PETRA in ns-3. The

evaluation process will include more network performance metrics: throughput, congestion window size, average delay, and jitters. Also, we will extend our study to test the fairness and the friendliness of the new implementation.

Chapter 6 summarizes the overall work in this study, and indicates the future work and recommendations.

