Vertical Handover Decision Making Using QoS Reputation and GM(1,1) Prediction

By David Giacomini

A Thesis

In The Department

of

Electrical and Computer Engineering

Presented in Partial Fulfillment of the Requirements

for the Degree of Master of Applied Science at

Concordia University

Montreal, Quebec, Canada

2012

CONCORDIA UNIVERSITY
SCHOOL OF GRADUATE STUDIES

This is to certify that the thesis prepared

By: David Giacomini

Entitled: "Vertical Handover Decision Making Using QoS Reputation and GM(1,1) Prediction"

and submitted in partial fulfillment of the requirements for the degree of

Master of Applied Science

Complies with the regulations of this University and meets the accepted standards with respect to originality and quality.

Signed by the final examining committee:

_____ Chair
        Dr. R. Raut

_____ Examiner, External
        Dr. C. Assi, CIISE   To the Program

_____ Examiner
        Dr. Y. R. Shayan

_____ Supervisor
        Dr. A. Agarwal

Approved by: _____
                        Dr. W. E. Lynch, Chair
            Department of Electrical and Computer Engineering

_____20_____   _____
                                                    Dr. Robin A. L. Drew
                                            Dean, Faculty of Engineering and
                                                    Computer Science

ABSTRACT

Vertical Handover Decision Making Using QoS Reputation and GM(1,1) Prediction

David Giacomini

Telecommunication consumers are fueling a demand for mobile devices that are rapidly increasing in their capability to provide a wider range of services. These services in turn are consuming more bandwidth and require richer quality of service (QoS) in order to ensure a good end user experience when performing activities such as streaming video content or facilitating voice over IP (VoIP). As a result, network providers are expanding and improving their coverage area while technology to establish Wi-Fi hotspots is becoming more accessible to every day users. This combination of increase in demand and accessibility, coupled with users' ever increasing expectations for high quality service presents a growing need to seamlessly optimize the use of the overlaid heterogeneous networks in urban areas to maximize the end user experience via the use of a vertical handover mechanism (VHO). Grey systems theory has been used in a wide range of systems including economic, financial, transportation, and military to accurately forecast time series based on limited information. In this thesis we build on a novel reputation based VHO decision rating system by proposing the use of the grey model first order one variable, GM(1,1), in the handover decision making progress. The low complexity of the GM(1,1) model allows for a quick and efficient prediction of the future reputation score for a given network, providing deeper insight into the current state of the target network. Furthermore, we analyze how this model helps balance the load across the heterogeneous networks employing its strategy.

Dedication

The journey that has been this thesis has been filled with many hurdles and challenges, none of which I would have been able to overcome with the support of numerous people. There are a few people though that I would like to express special acknowledgement and thanks to. I would like to thank my supervisor, Professor Anjali Agarwal, for her guidance and patience; my wife, Maryam Razmpoosh, for her love and encouragement; and my family, George, Robert, Manijeh, Bahram, and Mehrdad, for their constant support. Lastly, I would like to dedicate my efforts to my late mother, Sonia Giacomini, for all the love and support she unconditionally gave me throughout my life.

David Giacomini

Table of Contents

List of Figures

List of Tables

List Of Acronyms

Acronym	Definition
2G	Second generation wireless technology
3G	Third generation wireless technology
3GPP	Third Generation Partnership Project
4G	Fourth generation wireless technology
AGO	Accumulation Generation Operation
AHP	Analytical Hierarchical Process
AI	Artificial Intelligence
BE	Best Effort
BER	Bit Error Rate
BS	Basestation
CALA	Carribbean and Latin America Region
CDMA	Code Division Multiple Access
CDMA/CA	Carrier Sense Multiple Access Control with Collision Avoidance
CN	Core Network
CoS	Class of Service
CR	Consistency Ratio
DIA	Distance to the Ideal Alternative
ertPS	Extended Real-Time Polling Service
EURANE	Enhanced UMTS Radio Access Network Extension for NS-2
FIFO	First In First Out

GM(1,1)	Grey Model First Order One Variable
GSM	Global System for Mobile communications standard
HNN	Hopfield Neural Networks
IEEE	Institute of Electrical and Electronic Engineers
ITU	International Telecommunications Union
MAC	Model Algorithm Control
MADM	Multiple Attribute Decision Making
Mbps	Megabits per second
MEW	Multiplicative Exponent Weighting
MICS	Media Independent Command Service
MIES	Media Independent Event Service
MIH	Media Independent Handover
MIHF	Media Independent Handover Function
MIIS	Media Independent Information Service
MN	Mobile Node
MPEG	Moving Pictures Experts Group
NIST	National Institute of Standards and Technology
nrtPS	Non-Real-Time Polling Service
OFDM	Orthogonal Frequency Division Multiplexing
QoS	Quality of Service
RNC	Radio Network Controller
RSS	Receiver Signal Strength
rtPS	Real-Time Polling Service

SAW	Simple Additive Weighing
SINR	Signal to Interference and Noise Ratio
SMS	Short Message Service
TES	Transform Expand Sample
TOPSIS	Technique for Order Preference by Similarity to an Ideal Solution
UGS	Unsolicited Grant Service
UMTS	Universal Mobile Telecommunications System
UTRAN	UMTS Terrestrial Radio Access Network
VHO	Vertical Handover
VOIP	Voice Over IP
WCDMA	Wideband Code Division Multiple Access
Wi-Fi	Wireless Fidelity
WiMAX	Worldwide Interoperability for Microwave Access
WLAN	Wireless Local Area Network

1 Chapter 1: Introduction

The vertical handover mechanism required to facilitate seamless handovers in heterogeneous networks is a research topic that has seen significant activity since the advent of the IEEE 802.21Media Independent Handover standard [1]. Although this standard provides a framework for identifying and soliciting new heterogeneous networks, and the mechanism for making the actual physical handover, it does not describe the decision making algorithm that will ultimately trigger the handover to execute. In this thesis a novel vertical handover mechanism using a reputation based scoring technique is improved upon by leveraging a well-known prediction algorithm called the Grey Model First Order One Variable, also known as GM(1,1).

Before diving into these technical concepts it is important to understand some of the fundamental ideas that underlie them, such as the concept of end users being immersed in ubiquitous connectivity and the existence of overlapping heterogeneous networks. After clarifying this, the motivation behind this thesis and its application to real world scenarios is presented. This is then followed by a definition of the problem that the thesis is addressing and the list of challenges that were faced in tackling this problem. Subsequently, the objectives of the research are then stated in addition to the key contributions that were made as a result. Finally, the outline of the remainder of the thesis is discussed and a chapter summary is provided.

1.1 Ubiquitous Mobility

As technology in both end user handsets and network infrastructure increases, the ability for the consumer to be continuously connected is ever increasing. From an

individual's home Wi-Fi hotspot, to their 3G/Wi-Fi/WiMAX connected devices (mobile phone, tablet), to their workplace (company network and mobile devices), and everywhere in between, it's becoming more rare to be unable to instantly have access to someone or some service. In fact, finding a public telephone nowadays is something of a treasure hunt.

This perception of perpetually having access to the internet and its supporting telecommunication infrastructure is reinforced by numerous studies such as the 'Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2011–2016' [2]. This endeavor, which is part of a Cisco project to track and forecast the impact of visual networking applications on networks around the world, provides numerous facts about mobile traffic in the world; three pertinent findings in this report are as follows.

Man vs. Machine

The number of mobile connected devices will outnumber the human population by the end of 2012, and by 2016 there will be 1.4 mobile devices per capita [2].

Bandwidth Intensive Services

In 2011 mobile video traffic exceeded fifty percent of the global mobile traffic, sitting at fifty-two percent. By 2016, Cisco forecasts that two-thirds of the world's mobile traffic will be video [2].

Accessibility

In 2011 the average connection speed increased by approximately sixty-six percent, where the downstream speed was calculated as 315kbps. By 2014 an average consumer mobile connection will have a speed that surpasses 1 Mbps [2].

Given these measures, it is evident that society is moving in a direction of constant and complete connectivity. In order to meet this ever growing demand, it is important that network providers and device manufacturers effectively manage the heterogeneous network resources that are appearing within the environment. In doing so, not only will the customer be satisfied but it will lead to larger adoption and acceptance of the always connected mentality, driving more business.

1.2 Heterogeneous Networks

A mobile user who is streaming a music video on their smart phone while taking the bus to work will most likely leave the coverage area managed by one of its network provider's basestations and move into the next closest one. Although the user does not realize it, there is a handshake that occurs and the user's video streaming session is transferred from one basestation to the other, seamlessly. In this scenario, because the networks consist of the exact same technology, it is considered to be homogeneous. In the heterogeneous case, the consideration is to move from one type of network technology to another different one. The main difficulty here is that each technology behaves based on its own set of rules and languages; as a result, facilitating a common process across all these disparate technologies is no easy feat. The three main technologies leveraged in the experiments related to this thesis are WiMAX, Wi-Fi, and UMTS.

1.2.1 Wi-Fi

The IEEE 802.11 family of standards defines the behavior for the wireless local area network (WLAN) that is commonly referred to as Wi-Fi. The original incarnation of this standard came about in 1997 and since then numerous amendments have been layered on top of it. It is one of the de facto wireless standards of today's day and age, with a wide range of devices leveraging its technology to connect with the world. From mobile phones, to printers, to home appliances – the technology has become so accessible and affordable that more possibilities of how it can be leveraged in daily life continue to emerge. The Wi-Fi Alliance organization, a global non-profit trade association which promotes Wi-Fi, noted that currently one in ten people in the world use Wi-Fi [3].

The WLAN technology is primarily used in two ways. The first is as a private secure network used in indoor locations such as individual's homes or independent businesses. The other is as an open and free 'hotspot' made available by businesses such as cafes and restaurants, airports, and hotels to entice customers to do business with them. In North America, these networks usually consist of an access point, such as a wireless router, that is constantly broadcasting a signal on a 2.4 GHz frequency using Orthogonal Frequency Division Multiplexing (OFDM) [4]. All Wi-Fi equipped devices know to listen on this frequency and as a result, when they come into range of an access point, they can subsequently send and receive data via packets made up of Ethernet frames, using Carrier Sense Multiple Access Control with Collision Avoidance (CDMA/CA). As a result, the service is connectionless and contention based, which often results in mobile nodes located farther away from the access point being interrupted by mobile nodes situated closer. Typical Wi-Fi enabled devices in these scenarios support the common

802.11g standard and can manage data rates of about 54Mbps and a range of 30m indoors [4]. The latest version, 802.11n, promises up to three times the data rate and roughly double the range.

A third implementation of Wi-Fi is in the deployment of this technology across very large regions like densely populated metropolitan areas or large campuses. In these outdoor cases the range of technology can reach up to 300m but requires higher power transmitters that operate close to the threshold of what is permissible in the license-exempt frequency bands [4]. Furthermore, to achieve the widespread coverage a dense concentration of access points is needed in order to ensure that the coverage is ubiquitous [4]. Regardless of the range limitations, Wi-Fi is so widespread that it is a technology which must always be considered in VHO scenarios due to the practicality of its application in real life.

1.2.2 WiMAX

The IEEE 802.16 set of standards was originally completed in 2001. Although it is technically called 'Wireless Metropolitan Area Network' by IEEE, the WiMAX Forum, a non-profit industry led consortium which promotes the technology, commercialized it as 'Worldwide Interoperability for Microwave Access', WiMAX. The original standard has gone through several amendments, with the latest version being 802.16m-2011.

One of the main benefits of the WiMAX technology is its long range capabilities. A typical WiMAX basestation operating in the 2.3 GHz frequency provides a coverage range of about 1km [4]. Another key benefit of the technology is its ability to manage

quality of service (QoS). WiMAX has a connection-oriented design which was architected to support numerous types of applications, including those requiring high levels of QoS such as video streaming and conferencing, and voice over IP (VoIP). Specifically, these service levels are [5]:

- **Unsolicited Grant Service (UGS)**: Provides a maximum sustained rate, a maximum latency tolerance, jitter tolerance, and is primarily used for VoIP.

- **Real-Time Polling Service (rtPS)**: Provides a minimum reserved rate, a maximum sustained rate, latency tolerance, traffic priority, and is typically used for streaming audio or video.

- **Extended Real-Time Polling Service (ErtPS)**: Provides a minimum reserved rate, a maximum sustained rate, latency tolerance, jitter tolerance, traffic priority, and is typically used for VoIP.

- **Non-Real-Time Polling Service (nrtPS)**: Provides a minimum reserved rate, a maximum sustained rate, and traffic priority, and is typically used for file transfer protocol (FTP) traffic.

- **Best Effort (BE)**: provides a maximum sustained rate, traffic priority, and is typically used for all other traffic such as data transfer, web browsing, etc.

As a result of these benefits, WiMAX is used in several ways. One use of the technology is to provide a cheaper backhaul technology for 2G, 3G, and 4G networks in both developing countries and developed countries. Furthermore, the technology is used to provide last-mile broadband access to sparsely populated regions and internet access to large metropolitan areas, and can support peak rates of 63Mbps in the downlink and 28Mbps in the uplink [5].

Generally speaking, the number of WiMAX implementations globally is more numerous in Africa, CALA (Caribbean and Latin America Region) and Asia-Pacific, with some lesser presence in Europe, North America, and the Middle East, and the most popular frequencies of use are 3.5GHz and 2.5GHz [6]. Although there is some cross-over between the use of WiMAX vs. Wi-Fi in the same of metropolitan coverage, these two technologies can be seen as complimentary, with WiMAX providing wide access to numerous Wi-Fi implementations.

1.2.3 UMTS

In contrast to Wi-Fi and WiMAX, the Universal Mobile Telecommunications System (UMTS) is a third generation (3G) cellular mobile technology that was developed by the 3GPP (3rd Generation Partnership Project) and was based on the Global System for Mobile communications (GSM) standard. In fact, UMTS is one of the systems proposed under the International Telecommunication Union's (ITU) IMT-2000 Standard, and was the fruit of European and Japanese teams. The other main competitor to the UMTS implementation is the North American CDMA2000 system [7].

As UMTS is a full blown cellular mobile system, it consists of several key components, namely the mobile nodes (MNs), the UMTS basestations (BSs), the UMTS Terrestrial Radio Access Network (UTRAN), and the core network (CN). MNs communicate with the BSs by using Wideband Code Division Multiple Access (WCDMA) over the UTRAN. The UTRAN itself consists of multiple Radio Network Subsystems (RNSs) which in turn are made up of one or more Radio Network Controllers (RNCs); the RNC controls one or more BS. Finally, the CN connects the RNSs together [7].

Similarly to WiMAX, UMTS facilitates QoS management and supports four distinct classes of service [7]:

- **Conversational class**: The purpose of this class is mainly to service voice services such as VoIP. The typical delays in this service are targeted to be around 100ms or less; larger delays would result in poor end-user experiences.

- **Streaming class**: Typical usage of this class is for audio and video streaming where delays of larger than 100ms are acceptable due to buffering on the end-user's side. In contrast though, the bit error rate (BER) is typically lower in streaming than in conversational since users are usually more sensitive to the noise in music signals vs. voice conversations.

- **Interactive class**: This class is typically used for non-delay sensitive request / response applications, the most common being web browsing. In this class there still are upper limits on the tolerance of delay but it is usually on the order of seconds and not ms (e.g. how long it takes a web page to load).

- **Background class**: This class of service covers all applications where there is no delay sensitivity. Common examples include email and short message text (SMS).

The vast majority of UMTS implementations use the 1.9GHz to 2.025GHz and 2.11GHz to 2.2GHz frequency ranges. The frequency ranges are split among two types of channels: common channels and dedicated channels. Common channels are used for control purposes and thus consume a small amount of bandwidth, while the dedicated channels are used mainly to send the actual data on the uplink and downlink. The

maximum bit rate of the system is roughly 2Mbps and depends on the speed of the mobile user [7].

In general, UMTS is a relevant cellular mobile system to be considered in the analysis of this thesis as it represents a substantial user base around the global. Furthermore, the overlay of UMTS with WiMAX and Wi-Fi is not only technically feasible, but highly likely, as all three technology types are becoming more readily accessible to the general public.

1.3   Motivation and Application

The IEEE 802.21 Media Independent Handovers standard has been around for several years now and presents a very interesting framework to facilitate the handover between heterogeneous networks. With the continued increase in accessibility and decrease in cost to numerous types of network technologies, the question of how to leverage all the available network resources in urban environments is becoming a more relevant issue every day. As a result, one key area of exploration in the 802.21 standard is the actual decision making process itself; this section of the framework was left un-specified in order to let the industries which opt to utilize the framework determine how best to make the decision within their environment. As one looks into the topic though, the question of how to make the best decision is not so straightforward or simple. Considerations must be given to the complexity of the calculations that need to be made on the network or mobile node sides, the duration of the decision making process, and the effect of implementing the algorithm on the ecosystem of networks.

1.4 Problem Statement

There are a growing number of telecommunication consumers within large metropolitan environments which have access to one or more mobile devices (e.g. smartphones, tablets, and laptops) which can interface with numerous types of networks, such as 3G, Wi-Fi, and WiMAX. In addition, the applications offered by the mobile devices that are used by these consumers can often consume data which is latency sensitive.

Furthermore, given the ease of access to mobile content, consumers are actively utilizing these applications while travelling within their environment. The large cities in turn, are home to an abundant source of Wi-Fi hotspots and ubiquitous 3G coverage, and in some cases WiMAX coverage. As a result, the mobile device user is routinely entering zones of heterogeneous network overlap. Typical cases are seen in the following examples:

- A student streaming a music video while walking to class in the morning, and then stopping at a café for breakfast which hosts a free Wi-Fi hotspot.
- A professional talking with a friend via Voice Over IP on his/her mobile phone while taking the bus home from work, and crossing through one or more free Wi-Fi hotspot or a free WiMAX coverage area
- A teenager streaming a movie on their tablet while travelling with a parent in the car and crossing a large free WiMAX coverage area

When faced with this overlapping network situation, the mobile device should be smart enough to determine which of the available networks will offer the end user the

best quality of service, subsequently make the decision, and then seamlessly handover without the end user realizing that a change in network has taken place.

Given this, the purpose of this thesis is to provide an improved vertical handover decision making algorithm that will increase the quality of the decision being made by the mobile node in choosing the best network for the end user in terms of overall quality of service experienced. As a result, while the user is travelling through numerous areas of heterogeneous network overlap, there should typically be an increased QoS observed by the user during his/her trip in using the proposed algorithm.

1.5   List of Challenges

Enhancing the novel reputation based VHO algorithm presented by M. Zekri [8] and demonstrating that the technique was effective not only from the perception of a single mobile user but from the overall network load perspective itself was a task that presented numerous challenges.

Algorithm Complexity and Delay

One of the key aspects of the VHO decision making algorithm is the amount of time required to make the decision. A process which is too complex will result in a delay which will impact the end user's experience. Considering that the primary traffic leveraged in the experiments involves video streaming to model the typical scenario of people watching videos on their mobile devices while commuting, the question of delay is especially sensitive as it is readily noticeable in the application being used. As a result, any complexity required by the algorithm should be resolved if possible outside the decision making process itself.

Balancing the Network

The main focus of the thesis is how to leverage an optimized VHO decision making algorithm to facilitate the load balance of the network ecosystem while maintaining good QoS. In order to accomplish this, it is necessary to identify how the balancing of the network can be tested. In this way, a correlation must be done between the number of handovers that occur in the network and the perceived performance of the various networks with the ecosystem.

Network Simulation

Demonstrating that the proposed solution for the problem statement is correct requires the use of a network simulation tool. This aspect of the thesis proved to be one of the most challenging aspects due to the use of several modules created by other third parties, which were only compatible with an older version of the network simulation tool that was used, NS-2. As a result, not only was there a task of extending the open network simulation tool, but also in making the necessary modifications to ensure all the modules combined worked as expected.

1.6   Contributions of the Proposed Research

The majority of the work carried out in the literature surveyed regarding the VHO algorithm design focuses on the delay resulting from the process in addition to how it performs against other known methods. In this thesis, through building on an interesting approach, the impact against the network ecosystem is analyzed, providing a newer dimension of analysis in this space. Since the overall network performance is of great

concern to both the network providers and the end users of the network, this analysis is of value to a wide range of stakeholders.

In addition to analyzing the network environment, the proposed solution also leverages a well-known and used predictive algorithm, GM(1,1), a method that has not been seen in this space based on the articles surveyed at the time this thesis was written. This suggestion may open the door for other approaches to leverage predictive measures in their algorithms.

1.7   Thesis Outline

The remainder of the thesis is divided up into five main sections. The first section reviews the literature work that has been carried out in regards to vertical handovers, and focuses on the most popular approaches that have been covered recently.   This is followed by a review of the two main theories used to build the proposed solution, namely the reputation score-based model that facilitates quick decisions, and the GM(1,1) algorithm that provides the predictive behavior for time varying discrete sequences. Subsequently, the actual proposed solution is presented, discussing how the building blocks are gelled together and the reasoning behind other minor modifications made to form the overall solution. After this, the network simulation is presented, with details provided on the NS-2 simulator itself and its background, along with the outline of the network topology and configurations used in the thesis' experiments.   Once this is completed, the actual simulation performance results are discussed from the view point of one node and then many. In the first case, the behavior of a single node leveraging the proposed model vs. the original reputation model is analyzed. In the latter, the network as a whole is analyzed with each network type reviewed in conjunction with the mobile

nodes using the proposed and original decision making algorithms at various speeds. Finally, the overall findings and deductions are discussed in the conclusion.

1.8 Summary

In synopsis, the concept of numerous heterogeneous networks existing in an overlaid fashion within large metropolitan environments is becoming more and more common place across the world. With the advancement of technology and the increasing availability of cheap services and powerful devices capable of providing a wide range of applications, the problem of how to optimize this growing network ecosystem is fast becoming relevant. In this thesis, a novel reputation based vertical handover decision system is enhanced via the use of a prediction algorithm and the impact of this enhanced solution is measured against an individual mobile node and the ecosystem of networks, which consist of overlaid UMTS, WiMAX, and Wi-Fi coverages. The open source discrete event network simulator NS-2 is used to carry out the experiments.

2 Chapter 2: Literature Survey

2.1 IEEE 802.21 Standard

The purpose of the 802.21 Media Independent Handover Standard is to improve the user experience of mobile devices by providing a mechanism to seamlessly handover between networks, regardless of whether the networks are of the same technology type, regardless of whether the networks are from the 802 standards, and regardless of whether the networks are wired or not [1]. Due to the large number of technologies in existence today, and the high probability of the overlay of numerous network types within a dense urban environment, this standard provides an important tool in optimizing the end user experience in the mobile world. Numerous aspects of a vertical handover are covered by this Standard, as seen below.

Service Continuity

Service continuity implies that the service being consumed by the mobile device prior to the handover is maintained throughout and after the completion of the handover process, and resulting in minimal data loss. The quality of service may change from one network to the other due to network conditions and features, but user intervention should not be needed [1].

Quality of Service

The 802.21 standard addresses the QoS of the handover in two ways; one is in minimizing the actual delay and data loss resulting from the handover process itself through make before break handovers. The second aspect is in utilizing the QoS

information of the networks in the handover process, either through obtaining QoS information through the specific services in the standard to factor in the decision making process or in using services from the standard to set performance metric thresholds that trigger the handover process to occur [1].

Network Discovery

In order to make a handover, the mobile user must be able to identify new networks or realize that a network is within reach. This is made possible through specific services within the standard that users of the framework can invoke. Through network discovery, the mobile node can understand link availability, link quality, etc. [1].

Network Selection

Once one or more networks have been identified, the 802.21 standard provides the facility for the mobile node to obtain information about the networks in order to determine whether or not to make a handover and to whom. The standard does not define the handover policy itself, but makes available services to collect the necessary info in order to do so and to ultimately make the physical handover as well [1].

Power Management

One key benefit of the standard is that it does not require the mobile node to turn on a specific network technology interface in order to discover the associated network. As a result, the mobile node saves power by limiting the amount of time it needs to have multiple interfaces up in order to maintain a connection [1].

2.1.1 802.21 Structure

In order to facilitate the vertical handover, the Media Independent Handover (MIH) standard is built using the following key elements: the MIH Function (MIHF) and the MIH users. Their interaction is depicted in Figure 2.1:
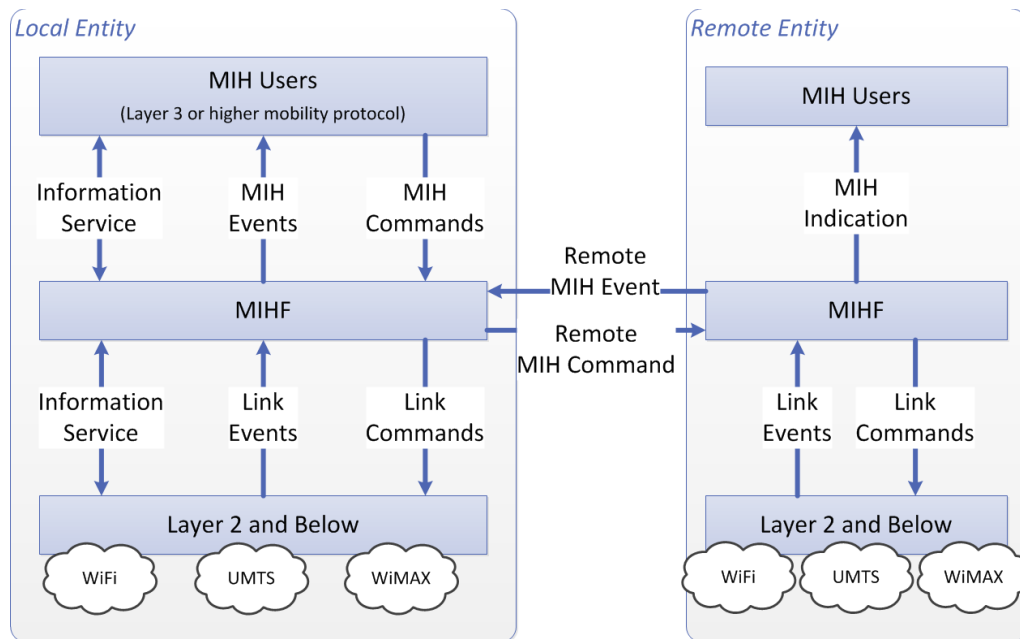
Figure 2.1: IEEE 802.21 structure

MIHF

As can be seen from Figure 2.1, the MIHF sits in between the MIH users, which are at the network layer 3 and above, and the network layer 2. The purpose of the MIHF is to provide abstract services to the MIH users through a generic service access point that is media independent, and to also retrieve information from the layer 2 technologies through service access points that are implemented for specific media types [9]. The MIHF can also support communication of events and commands with remote MIHFs, and equipment that is implemented to leverage this technology should be compatible with legacy equipment [1].

MIH User

Where the MIHF is the provider of abstract services, the MIH user is the consumer of these services, typically represented by the mobile device or mobility management application that would leverage the services to enable the handovers to occur. MIH users can subscribe to specific services to be notified about critical events that could result in them triggering the handover process, or they can use services to perform actions like initiating the handover process, collecting information about certain links, or even setting thresholds for connections to facilitate QoS [9].

2.1.2 802.21 Services

There are three types of service that are offered and managed by the MIHF and which are consumed by the MIH users: the event service, command service, and information service.

Event Service

The purpose of the Media Independent Event Service (MIES) is to provide a mechanism for reporting on changes that occur at the link level, such as the status of a given link (e.g. Link_Up) or the quality of the links performance (e.g. Link_Parameters_Report). The events can also provide some advance notice of changes occurring, e.g. Link_Going_Up or Link_Going_Down. In general, the events are considered as discrete and can be split into two categories: Link Events and MIH Events.

Link Events originate in the link layer (e.g. in the WiMAX link) and propagate to the MIHF. From there, the Link Event can continue propagating as is, or with some

additional processing by the MIHF, to the MIH users that have subscribed to that particular type of event. MIH Events on the other hand are events that arise from the MIHF or are link events that the MIHF chose to forward to the MIH user. An event being communicated from one MIHF to another MIHF is considered as a remote event. As a result, all the Link Events are considered to be local since they only go from the link layer to the MIHF or the local MIH users.

Command Service

The Media Independent Command Service (MICS) provides a facility for the MIH user to perform several kinds of actions. Firstly, this service can be used by the MIH user to request link parameters for specific links (e.g. what is the throughput, delay, etc). Secondly, the service is used by the same audience to configure thresholds on specific links in order to facilitate handovers and maintain QoS. It is worth noting that as a result of these first two types of services, corresponding events will typically be triggered in response (e.g. a threshold being breached on the configured link). A third type of command is the set surrounding the actual handover process, e.g. querying for the network (MIH_MN_HO_Candidate_Query) or committing to handover to that network (MIH_MN_HO_Commit) [1].

Similarly to the event service, Commands can be either local or remote. Local events originate from the MIH user and propagate to the MIH user's corresponding MIHF. Remote events are sent by the MIH user to the local MIHF who then forwards this request to the peer MIHF. The MIH commands received by the MIHF are typically

converted into link commands and sent to the link layer to carry out the original request in the MIH command [1].

Information Service

The purpose of the media independent information service (MIIS) is to build a global view of the heterogeneous ecosystem so that the MIH user or MIHF can leverage the collected information to help optimize the handover decision process. Typically the information captured by this service is static network information, such as whether security or QoS is supported by the network, or availability of networks in the geographical vicinity. The latter can help facilitate a vertical handover situation by indicating to the mobile node an appropriate place to move towards in order to trigger the handover [1].

2.1.3 802.21 Process

Given the three sets of services, the typical sequence of service instantiations in a mobile node initiated vertical handover process can be seen in Figure 2.2 [10]. In step 1, the mobile node requests network related information from the MIIS via the MIHF. The request is received by the MIIS and a response is generated with the appropriate information and sent to the MIHF, who then forwards the response to the MIH user in messages 2 through 4. From the message the mobile node, which is currently residing in a UMTS network, understands that a WiMAX network is currently in the vicinity. As a result, it turns on its WiMAX interface to determine when it has reached the coverage point. Once detected, the interface sends a link detect event to the MIHF who in turn notifies the MIH user in message 6. Based on this, the MIH user requests up to date

information from the target network via the current network point of service, such as what the available QoS support is and resource availability. Based on this response, the mobile node determines that it wishes to make a handover to the target network, as indicated in message 12. Once the network layer-2 connection is made to the new network, the mobile node's corresponding WiMAX interface sends a handover complete message to the MIHF in message 14 which is then forwarded to the MIH user. At this point the higher layer handovers can complete their process; the MIH user signals this completion in message 16, and the new network confirms it in messages 17 through 20.
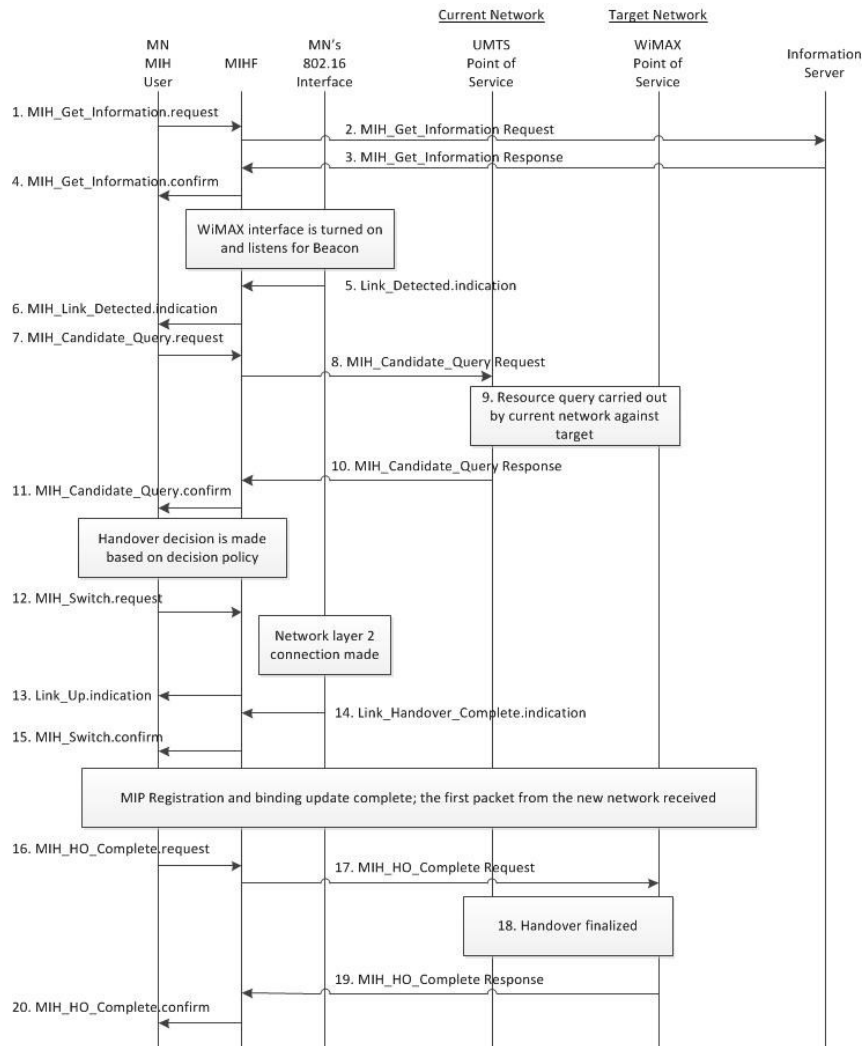
Figure 2.2: IEEE 802.21 handover protocol sequence

2.2   Vertical Handover Decision Models

The proposed model presented in this thesis leverages a decision model based on a reputation scheme employed in the heterogeneous networks, but in the research area of vertical handovers there are numerous types of models that are being employed and suggested. In this section the main algorithms are presented, and their strengths and weaknesses are discussed.

2.2.1   Network Condition Based Decisions

In order to minimize the delay incurred by the vertical handover process it is important to design an algorithm or decision policy that is as simple as possible. In this light, numerous studies into the VHO question have leveraged simple measurements of the network condition to design a decision algorithm. A typical example of this is through using the receiver-signal strength (RSS) of the mobile node to affect the handover. The assumption here is that the stronger the receiver signal, the closer the mobile node is to the access point of the network, and the higher the chance is that it will receive a better quality of service. In [11], the authors investigate the RSS technique and determine that although it provides a straightforward basis for making a handover decision, it is not the most efficient method since it results in higher power consumption due to the need to keep multiple interfaces of the mobile node on in order to sense the RSS in the current and target networks. It is commonplace to find studies in the VHO space measuring the performance of a proposed VHO model against the RSS approach to gauge success.

An alternative to the RSS approach is the use of the Signal to Interference and Noise Ratio (SINR), which is a function of the signal the mobile node experiences, based on its location in relation to the network access point, and the associated interference in the network. In [12] this approach is taken and compared against the standard RSS technique under various data rates applied to the network. The result is that the SINR provides a higher overall throughput experienced by the mobile device at all the data rates, and that this increase is more predominant at higher data rates.

A third method of utilizing the network condition as criteria in making the VHO decision, is through the use of the available data rates in the networks, as documented in [13]. In the aforementioned paper, the proposed model is implemented through the mobile node leveraging the available MIH MICS and MIES services to obtain the current and target network data rates. Based on this information, the mobile node chooses the network with the highest available data rate. In order to measure the benefits of this algorithm, the authors compare the proposed model against the RSS model, and the result is that the data rate method provides higher data rates for the mobile node when it takes the RSS into account.

Ultimately, although the techniques based on the network condition provide a simple mechanism for making the decision that incur low delay, they typically do not provide the best overall throughput and QoS for the mobile device. As a result, if the mobile node is leveraging services or applications that have high QoS requirements, this technique is not best suited to facilitate the handovers.

2.2.2 Multiple Attribute Decisions

The Multiple Attribute Decision Making (MADM) model attempts to optimize the decision making process through the use of several inputs that are typically based on the network condition and performance. In [14] the authors use two MADAM algorithms to attack the VHO problem, namely an Analytic Hierarchical Process (AHP) to assign weights against the alternative networks in relation to the current network [14], and then use Technique for Order Preference by Similarity to an Ideal Solution (TOPSIS) in order to select the best option [14]. The criteria used in this approach are the network types and several network performance metrics such as cost per byte, total bandwidth, allotted bandwidth, utilization, delay, jitter, and packet loss. The model is then compared against standard implementations of MADM, which only use TOPSIS or Distance to the Ideal Alternative (DIA) [14], to identify the improvement gained through the proposal.

The authors in [15] also leverage MADM, using AHP for the weight assignment of specific traffic parameters and Simple Additive Weighing (SAW) and Multiplicative Exponent Weighting (MEW) to calculate the score which represents the overall QoS offered by the network. Through comparison with the basic RSS approach, it is shown that the MADM model performs better through lower packet dropping rates and smaller handover times.

In general, the MADM model is seen as a higher complexity model as compared to the network condition approach. Although it provides more QoS awareness in the decision making algorithm through the use of network performance metrics, it typically results in higher handover delays due to the complexity of the decision making computation [15].

2.2.3 Artificial Intelligence Decisions

The most popular implementations of vertical handover decision algorithms using artificial intelligence (AI) are through the use of fuzzy logic or neural networks. In order to utilize these methods effectively, they are typically paired with other existing VHO decision mechanisms in order to build more intelligence around using the knowledge of the system in conjunction with the calculated or obtained information from the VHO algorithms.

In [16], the authors use fuzzy logic in association with a TOPSIS MADM model. In order to make a decision which optimizes the QoS experienced by the user, the TOPSIS model, uses resource availability, RSS, mobile node speed, network type, network link business cost, and security as criteria in the analysis. Since these performance metrics vary significantly across network types and even within homogeneous types, multiple fuzzy logic controls are applied in order to better rationalize the gathered information. The success of the proposed scheme is assessed based on its performance against several other methods including the RSS approach, where it was shown that the fuzzy logic based approach results in fewer handovers, lower handover failure rate, and a higher percentage of users which are assigned their preferred network or lower cost networks.

Alternatively to the fuzzy logic approach, the authors in [17] propose a deviation to the Hopfield Neural Networks (HNN) mechanism in order to optimize the VHO decision, since the original HNN method is considered to be a powerful optimization tool for complex problems. In this model, numerous network performance metrics such as application bit rate, application delay, application type, current/target network capacity,

current/target network delay, mobile node speed, handover imminent trigger, and SINR of the point of attachment link are assessed and mapped to five key calculated metrics that are used in the actual HNN decision process. The five metrics are rate, delay, application type, velocity, and handover imminent, where the latter is derived based on the severe deterioration of the network performance. When the proposed model is compared against the typical RSS model, it is clear that there are improvements in the five metrics being evaluated.

Similarly to the MADM model, there are significant benefits in using this model to choose the network with the best QoS available, but the cost is typically in higher delays in the actual handover process. The fuzzy logic and neural network algorithms are usually considered as highly complex calculations, resulting in this delay [15]. As such, if the mobile node is using delay and jitter sensitive applications there is some risk involved in leveraging this mechanism.

2.3 Summary

A key driver of the vertical handover problem space is the advent of the 802.21 media independent handover standard. Through the definition and adoption of a generic multi-technology handover protocol, it is more likely that mobile devices and network operators will support this seamless transition of service between distinct network technologies. Although the standard provides a framework for discovering and identifying networks, and a protocol to facilitate the actual handover, it does not provide any insight into how best to implement a decision policy algorithm either in the mobile node or in the network itself. As a result, numerous methods and studies have been carried out to fill in this gap, with the most popular mechanisms including the use of

network conditions such as mobile node RSS and SINR, MADM algorithms leveraging theories such as AHP, TOPSIS, SAW, and MEW, and finally AI based calculations utilizing concepts such as fuzzy logic and neural networks.

The majority of the findings have found that within the three, the simplest is the network condition based method, yielding low delays but not realizing a strong understanding of the QoS offered by the available network set. The MADM and AI approaches on the other hand, offer significantly optimized methods of choosing the network with the best offered services, but results in yielding typically larger handover delays due to the complexity of the algorithms [15]. As a result, there is still much research activity in this space to identify the best well rounded mechanism.

3 Chapter 3: Building Blocks

The proposed vertical handover decision algorithm presented in this thesis is built using two key components. The reputation model originally proposed in [8] forms the basis of the proposed model and provides a framework for enabling quick VHO decisions within the heterogeneous ecosystem. Secondly, the mathematical model devised by Deng, in [18], provides a mechanism for making predictions based on a limited set of previous sequential data inputs. These two building blocks are further expanded upon in this section.

3.1 Reputation Model

Facilitating a quick vertical handover decision algorithm is of the utmost importance when implementing the 802.21 media independent handover standard, especially when the user is consuming delay sensitive information. In order for the handover process to be seamless, the user must not be aware of the events happening in the background. In light of this, one solution to resolving the problem of delay is making the actual decision simple. M. Zekri et al, in [8], achieved this by offloading much of the decision making complexity into the time period outside of the decision making process itself, through the use of a novel scheme employing agents residing in the network and on the individual mobile nodes.

3.1.1 Reputation Model Structure

The essential premise of the reputation model is to leverage special agents whose role is to compute and aggregate a score which provides a numerical representation of the quality of service currently offered by a specific network at a given point in time. As a

result, when a mobile node discovers a new network, it simply needs to obtain the latest representative score of the network, compare it to its current network's score and an acceptable threshold, and then make the decision. To this end, there are two specific types of agents which provide distinct services within this reputation framework: the mobile reputation agent and the network reputation agent.

Mobile Reputation Agent

A mobile node which leverages the 802.21 media independent handover standard must be physically capable of supporting multiple types of network technologies. As a result, such a mobile node will have an interface to each of the networks it supports. Similarly, to implement the reputation model, a mobile node reputation agent should exist on each of the mobile node's interfaces.

The purpose of these agents is to then collect real-time performance metrics upon entering and leaving the respective network in order to calculate sample reputation scores. The reputation score is then converted into a binary trust system where the network is evaluated as either good or bad. More specifically, let N be the set of available networks, and M_n the set of mobile nodes that are currently connected to network n, where $n \in N$. The sample score R is then calculated by each mobile node in the network, $r(m, n)$, where $r \in R$ and $m \in M_n$. When the mobile node enters or leaves a specific network, only the reputation agent responsible for this network collects statistics and computes the score, while the other reputation agents are dormant waiting for their network to be accessed.

In addition, when the mobile node reaches the coverage area of a new network, the mobile reputation agent responsible for the network technology type is leveraged to communicate with the target network residing agent to obtain the advertised score. Once obtained, the mobile node has the necessary information to make a decision regarding whether or not to handover to the new network.

Network Reputation Agent

Each distinct network within the network ecosystem leveraging the reputation model requires one network residing reputation agent. The role of this agent is two-fold − one is to continuously collect all the real-time binary converted sample scores calculated by the mobile residing reputation agents currently utilizing the network and subsequently form the running reputation score of the network, $R_n = \{r(m_i, n)|m_i \in M_n\}$. Since it is a binary reputation system, the sample score perceived as good is represented by $r^+(m, n) = 1$, and $r^-(m, n) = -1$ when it is considered bad. This aggregated score synthesizes what the overall quality of service offered by the network is at the time of calculation.

The second purpose of the network residing reputation agent is to then provide this aggregate score upon request to any mobile residing reputation agent requesting it. As a result, the network agent advertises the network's QoS to facilitate handovers by mobile nodes entering within its vicinity.

3.1.2 Calculating the Score

The principle question that comes to mind when analyzing this reputation system is how does one determine whether the score is good or bad? The execution of this

decision within [8] is carried out by the mobile node reputation agent when the MN initially enters and leaves the network. As a result, although determining the answer to this question may be complicated, it is completed outside of the decision making process itself and does add delay to the handover process.

In order to facilitate this choice, the calculated score, Q_n, must be compared against a threshold score, Q_{th}, which is used as a basis of comparison. The actual and threshold scores themselves are built using the key performance metrics used to quantitatively measure the quality of service levels of a network, specifically: bit error rate (ber), delay (del), jitter (jit), and bandwidth (bw). Depending on the class of service (CoS) in question, the Q_{th} is calculated accordingly, since the minimum requirements of each CoS vary greatly and put importance on different metrics within the list. As such, since the Q_{th} is representative of a CoS and does not change, its value can be calculated once beforehand.

In building the score, a specific weight is assigned to each QoS performance metric depending on the applicable CoS in question through the use of the AHP. Specifically, each QoS metric is correlated to a goal of a CoS; subsequently, through the use of 9 different importance levels, the goals are prioritized in relation to each other within a CoS. Assume p_{ij} is the prioritization of Goal i, G_i, in comparison to Goal j, G_j, where $\{i, j\} \in \{ber, del, jit, bw\}$, then, as seen in [8]:

- $p_{ij} = 1$ when the two goals are equal in priority

- $p_{ij} = 3$ when G_i is weakly more important than G_j

- $p_{ij} = 5$ when G_i is strongly more important than G_j

- $p_{ij} = 7$ when G_i is very strongly more important than G_j

- $p_{ij} = 9$ when G_i is absolutely more important than G_j

Based on the above comparisons, the AHP matrix is established and normalized, as can be seen generically across any CoS in Table 3.1.

Table 3.1: Normalized AHP matrix generalized for a class of service [8]

CoS$_i$	BER	Delay	Jitter	BW
BER	1	p_{12}	p_{13}	p_{14}
Delay	$1/p_{12}$	1	p_{23}	p_{24}
Jitter	$1/p_{13}$	$1/p_{23}$	1	p_{34}
BW	$1/p_{14}$	$1/p_{24}$	$1/p_{34}$	1

The normalized values, p_{ij}, are then used in [8] to calculate the weight, W_i, associated to a QoS parameter in a given CoS as:

$$W_i = \frac{p_{i1} + p_{i2} + p_{i3} + p_{i4}}{4} \tag{3.1}$$

Once the weights have been defined, the additional preparatory step required for calculation of the scores is to define normalization factors, X_{min} or X_{max}, for each of the performance metrics and within each CoS in order to ensure that the inherent value of the QoS parameter does not drive the score. This can be readily seen when considering that the raw value of jitter is typically in the order of one thousandth whereas bandwidth can be on the order of one million. The normalizing factor will then be applied to the raw metric, X_{raw}, based on the interpreted worth or cost of the parameter in order to obtain X_{norm}, the normalized value. As discussed in [8], if the parameter is valued more the lower it becomes, as is the case for jitter, delay, and bit error rate, then the normalizing

equation can be seen by (3.2). Otherwise, if the higher value has more worth, as in the case of bandwidth, the normalizing equation is described by (3.3).

$$X_{norm} = \frac{X_{min}}{X_{raw}} \tag{3.2}$$

$$X_{norm} = \frac{X_{raw}}{X_{max}} \tag{3.3}$$

Now, with both the weights and the normalizing factors defined for each QoS performance metric, the Q_{th} for a given CoS, c_i, can then be calculated (as demonstrated in [8]) by normalizing the threshold value of the metric and CoS in question and applying the corresponding weight, as seen in (3.4):

$$Q_{th}(c_i) = W_{ber(c_i)} * \frac{ber_{min}(c_i)}{ber_{th}(c_i)} + W_{del(c_i)} * \frac{del_{min}(c_i)}{del_{th}(c_i)}$$

$$+ W_{jit(c_i)} * \frac{jit_{min}(c_i)}{jit_{th}(c_i)} + W_{bw(c_i)} * \frac{bw_{th}(c_i)}{bw_{max}(c_i)} \tag{3.4}$$

Similarly, in [8], the actual sample reputation score, $Q_n(c_i)$, is calculated by the mobile node upon entering or leaving a given network in the same fashion as the Q_{th}. The only difference here being that the sample performance metric obtained by the mobile node is normalized instead of the threshold value, as can be seen in (3.5):

$$Q_n(c_i) = W_{ber(c_i)} * \frac{ber_{min}(c_i)}{ber_n} + W_{del(c_i)} * \frac{del_{min}(c_i)}{del_n}$$

$$+ W_{jit(c_i)} * \frac{jit_{min}(c_i)}{jit_n} + W_{bw(c_i)} * \frac{bw_n}{bw_{max}(c_i)} \tag{3.5}$$

Once the sample reputation score is calculated, it is compared against the threshold value obtained in (3.4) and ranked as either good or bad. The ranked score is then sent to the corresponding network residing agent for aggregation.

3.1.3 Aggregating the Score

Once the sample score is received by the network residing agent, the agent begins the aggregation process to determine what the running overall network score is. It is important to note that in [8], the network agent is modeled as processing multiple received scores within a given interval. Furthermore, the network residing agent can allocate a separate weight for favorable vs. unfavorable scores, w+ or w-, as seen in (3.6), in order to give more importance to poor network behavior since this is typically of the most concern [8].

$$r_{sample}(t) = w^+ \sum r^+(m, n) + w^- \sum r^-(m, n) \tag{3.6}$$

In addition to putting emphasis on the negative scores, the network residing agent also places greater importance on the recently received scores in [8] via the use of a discounting factor, $\gamma \in [0, 1]$. This discount is applied against the summed and weighted new score in (3.6) and then combined with the current network score to form the running aggregate score as seen in (3.7).

$$r_{aggr}(t) = \begin{cases} r_{sample}(t) & , \ t = 1 \\ (1 - \gamma) * r_{aggr}(t) + \gamma * r_{sample}(t), & t \geq 2 \end{cases} \tag{3.7}$$

The running aggregate score $r_{aggr}(t)$ is then provided by the network agent upon request from other mobile agents, to facilitate handover.

3.1.4 Reputation VHO Process

The reputation based vertical handover model proposed in [8] focuses mainly on the scenario where a mobile node is within a given network and while communicating in the network identifies that another distinct network is within reach. In fact, this speaks to the alternative handover portion of the overall model proposed, as seen in Figure 3.1, where the remaining scenario in the model is perceived as an imperative handover.
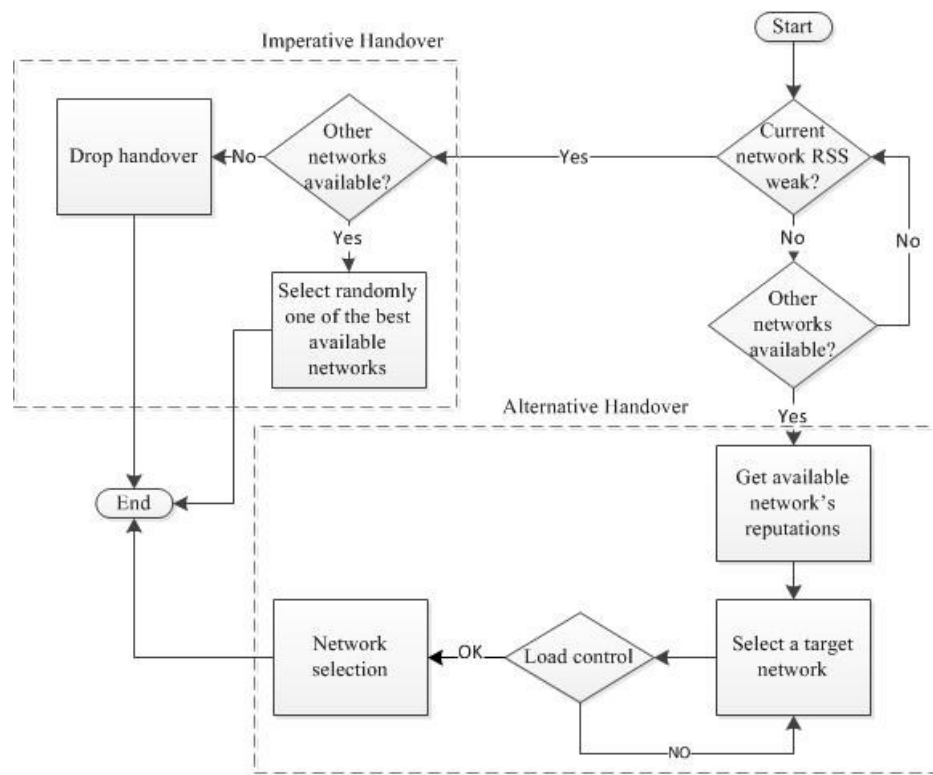
Figure 3.1: Vertical handover decision algorithm process flow proposed in [8]

The imperative handover module is provided to account for the situation where the mobile node exits the coverage area of the current network and is forced to choose somewhere to handover to in order to maintain the connection. Alternatively, this scenario could also be due to the mobile node physically losing connection to the current network for a variety of reasons, from its own interface malfunctioning to the network

provider having issues. Either way, when faced with this situation the mobile node will randomly choose from the list of available known good networks or if none are found, drop the connection.

3.2 Grey First Order One Variable Prediction

Grey system theory has been widely used across many disciplines of study and industries of business for decades, and as a result it has been a very active topic of research with many extensions and enhancements made to the models. Consequently, it is important to be aware of the various ways the grey system model can be leveraged in order to best utilize it in the scenario in question. In this section, the typical variations of the GM model are discussed, and the applications and derivation of the GM(1,1) are presented in order to provide more context into one of the key building blocks of the proposed model.

3.2.1 Grey Systems

Grey system theory-based models were first introduced by Deng in [18] in 1980, and have been used for predicting time series functions in a multitude of areas ever since due to its practicality and efficiency in estimating the behavior of unknown systems with a small data set [19]. The data set is always based on the most recent values from a time varying sequence, which are always positive and sampled at a fixed frequency. Since the initial creation of grey systems theory numerous modifications and extensions have been made.

GM(n,m) Model

At the most generic level, a *GM(n,m)* model analyzes an unknown system through an nth order differential equation with *m* variables. The most popular implementation of this model is the GM(1,1) version, which will be expanded upon in further sections. In addition to this, there has also been research in GM(2,1) models as seen in [20] and in GM(1,N), as seen in [21].

GM(1,1) Model

The simplest form of the grey system is the first order one variable implementation. In this form of the model, the coefficients of the associated differential equations are time varying, which leads to the model being re-defined every time a new input is received. When the data set used to calculate one predicted value is then shifted and used in conjunction with the latest value, the model is referred to as GM(1,1) rolling. In either case, in order to smooth out the incoming random data the model applies an operator called the Accumulating Generation Operator, which essentially sums the data iteratively [19]. This will be covered in greater detail in the derivation section (3.2.4).

The Grey Verhulst Model

The main purpose of the Verhulst model is to predict systems that contain saturation points, where the response is similar to an S curve with initial and end growth regions moving slowly and joined via a rapid intermediate section [22]. Examples of applications of this model include population growth estimations [22] and prediction of business operation cash flows [23].

Grey Residual Error Correction Models

For long sequences that tend to be more error prone using the GM(1,1) model, residual error correction methods have been employed to increase the performance of the first order one variable system. This has led to numerous implementations of such systems, as seen in [19] through the use of Fourier series, in [24] using a back-propagation neural network, and in [25] through the use of a model algorithm control (MAC).

3.2.2 Applications of GM(1,1)

The simplicity of the GM(1,1) model lends itself to being applied in numerous ways and scenarios. In [19], the first order one variable model is compared against the other typical variations of the GM model in the ability to forecast foreign exchange rate behavior. This analysis demonstrates that in non-saturation like conditions the GM(1,1) tends to perform better than the other variations of the model. Furthermore, it demonstrates an important use from a business perspective in forecasting key economic data. The business applications can also be seen in [26], where the authors modify the standard GM(1,1) model to form a discrete version of it in order to forecast real estate prices of a given area.

From a government perspective, the authors in [27] use the model to track the progression of a country meeting its greenhouse gas emission reduction commitments. This enables the group to demonstrate that the model can be utilized by policy makers to quickly understand the short term results of policies put into effect. In [28], the model is applied to forecast the amount of revenue received by tourism.

Drawing closer parallels to the work done in this thesis, the GM(1,1) model has also been used in monitoring traffic within a given homogenous network, as seen in [29] and [30], but this has typically been done through the combination of the GM model with a neural network.

3.2.3 Derivation of GM(1,1)

The first order one variable implementation of the grey system must have an input sequence that contains only positive values. In order to model this, let $X^{(0)}$ represent a time series with n values which is to be analyzed for prediction,

$$\mathbf{X}^{(0)} = (\mathbf{x}^{(0)}(1),\ \mathbf{x}^{(0)}(2), \dots,\ \mathbf{x}^{(0)}(\mathbf{n})),\ \mathbf{n} \geq \mathbf{4}. \tag{3.8}$$

This time sequence is then applied to an Accumulation Generation Operation (AGO) function in order to build $X^{(1)}$ and smoothen the randomness of the original values. This new sequence can be observed in (3.10) to be constantly growing,

$$\mathbf{X}^{(1)} = (\mathbf{x}^{(1)}(1),\ \mathbf{x}^{(1)}(2), \dots,\ \mathbf{x}^{(1)}(\mathbf{n})),\ \mathbf{n} \geq \mathbf{4}, \tag{3.9}$$

$$\mathbf{x}^{(1)}(\mathbf{k}) = \mathbf{AGO} * \mathbf{X}^{(0)} = \textstyle\sum_{\mathbf{i}=\mathbf{1}}^{\mathbf{k}} \mathbf{x}^{(0)}(\mathbf{i}),\ \mathbf{k} = \mathbf{1}, \mathbf{2}, \mathbf{3}, \dots, \mathbf{n}. \tag{3.10}$$

Subsequently, the AGO generated sequence is then used to define a mean sequence of adjacent data, $Z^{(1)}$, as follows:

$$\mathbf{Z}^{(1)} = (\mathbf{z}^{(1)}(1),\ \mathbf{z}^{(1)}(2), \dots,\ \mathbf{z}^{(1)}(\mathbf{n})),\ \mathbf{n} \geq \mathbf{4}, \tag{3.11}$$

$$\mathbf{z}^{(1)}(\mathbf{k}) = \tfrac{1}{2}\mathbf{x}^{(1)}(\mathbf{k}) + \tfrac{1}{2}\mathbf{x}^{(1)}(\mathbf{k}-\mathbf{1}),\ \mathbf{k} = \mathbf{2}, \mathbf{3}, \dots, \mathbf{n}. \tag{3.12}$$

From Deng's work in [18], it can be shown that the AGO generated sequence can be modeled by the first order differential equation (also known as the whitening equation) in

$$\frac{dx^{(1)}(t)}{dt} + ax^{(1)}(t) = b, \tag{3.13}$$

where a and b are referred to as the development coefficient and grey input respectively. Intuitively, from (3.10) one can also deduce that

$$\frac{dx^{(1)}(t)}{dt} = x^{(1)}(k) - x^{(1)}(k-1) = x^{(0)}(k). \tag{3.14}$$

As a result, by substituting (3.10), (3.12) and (3.14) into (3.13), one can obtain the Grey Differential Equation:

$$x^{(0)}(k) + az^{(1)}(k) = b. \tag{3.15}$$

In order to solve equation (3.15), one must obtain the solution for both the a and b parameters. This can be achieved through the use of the Least Square Error Method as follows:

$$[a, b]^{T} = (B^{T}B)^{-1}B^{T}Y, \tag{3.16}$$

where

$$Y = [x^{(0)}(2), x^{(0)}(3), ..., x^{(0)}(n)]^{T}, \tag{3.17}$$

$$B = \begin{bmatrix} -z^{(1)}(2) & 1 \\ -z^{(1)}(3) & 1 \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ -z^{(1)}(n) & 1 \end{bmatrix}. \tag{3.18}$$

Once *[a, b]* is solved, and knowing that the initial condition is $x^{(1)}(0) = x^{(0)}(1)$, the solution to the first order differential equation is:

$$x_p^{(1)}(k+1) = \left[x^{(0)}(1) - \frac{b}{a}\right]e^{-ak} + \frac{b}{a}, \tag{3.19}$$

where $x_p^{(1)}$ is the AGO generated value at the predicted time k+1. As such, in order to determine the actual predicted value at k+1, the inverse AGO is applied against (3.19) in order to obtain

$$x_p^{(0)}(k+1) = \left[x^{(0)}(1) - \frac{b}{a}\right]e^{-ak}(1 - e^a). \tag{3.20}$$

Additionally, the solution defined by (3.20) can be expanded to obtain a predicted value at time (k + H),

$$x_p^{(0)}(k+H) = \left[x^{(0)}(1) - \frac{b}{a}\right]e^{-a(k+H-1)}(1 - e^a). \tag{3.21}$$

3.3 Summary

The proposed vertical handover decision making algorithm presented in this thesis is composed of two key theories. The novel reputation based scheme proposed in [8] utilizes a mobile node residing agent which calculates a score upon entry and exit of a given network technology which quantifies the quality of service offered by the network. This score is communicated to the network residing agent, which aggregates the score

and then advertises it upon request to other mobile nodes interested in joining the network. As a result, the complex calculations needed to compute the score are performed outside the decision making process, and the decision itself is left as a simple comparison. The GM(1,1) algorithm on the other hand is the most popular implementation of the grey system proposed in [18]. Through the use of only a few sequential positive data points the algorithm can be used to predict the next future value. Due to the simplicity of the model, the theory has been used in countless applications.

4   Chapter 4: Proposed Model

Numerous approaches have been reviewed in the pursuit of facilitating a vertical handover decision module that results in not only minimal delay during the handover process but in making a choice that provides the most benefit to the applications currently in use by the mobile device.   Typically, optimizing one of the two areas, such as delay incurred from the decision process, tends to lead to diminished benefits in the other, i.e. choice of network based on provided service, and vice versa.   In this way, the approach proposed in [8] is very interesting since it manages to keep the delay resulting from the decision algorithm low due to its simplicity, while ensuring that the decision factors in the key QoS metrics to optimize the user experience.   In order to build upon this model and improve it, the primary goal was to improve the optimization of the choice of network without impacting the speed in which the decision is made.

4.1   Predicted Network Reputation

One key aspect of the reputation model proposed in [8] is that the authors assume that the reputation agent residing on the network, who's role it is to aggregate the calculated sample scores, as displayed in Figure 4.1, will receive multiple sample scores within a given time frame.   Furthermore, the actual sample score that is calculated by the reputation agent residing in each mobile node upon entry into and exit from the network, is converted to a binary score of -1 or 1.
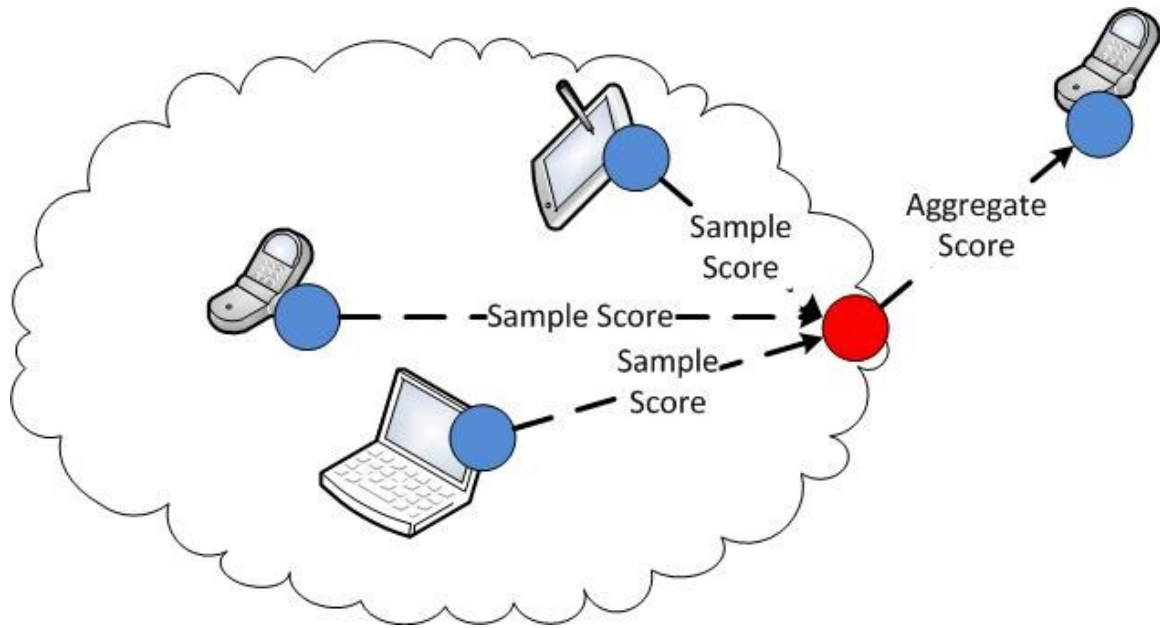
Figure 4.1: Interactions between the network (red) and mobile node (blue) residing reputation agents

As a result, instead of aggregating a score that is rich in QoS information, the network reputation agent is aggregating an amount which really represents the amount of mobile nodes experiencing a better than threshold service (represented by $\sum r^+(m,n)$) in comparison to mobile users who are observing a below threshold experience (represented by $\sum r^-(m,n)$), as see in equations (3.6) and (3.7).

This in itself does indicate some degree of quality of service information, since it infers that if the number is more positive, there are more nodes experiencing better than threshold service. But what this does not indicate is how good the service is. Specifically, one could argue that based on this theory if a large number of nodes are experiencing slightly higher than threshold QoS in a given network A, and a small amount of nodes are below the threshold in their observed service, then network A would be rated higher than a separate network B where a much lower number of nodes were

experiencing an incredibly high QoS and none were experiencing under threshold service.

As a mobile user streaming delay and jitter sensitive data, in order to optimize the end user experience, the goal would be to reduce the aforementioned QoS metrics as much as possible. In this way, the higher the QoS offered by a given network, the more advantageous it would be seen by the user consuming the time sensitive information. This would imply that the preferred network in the hypothetical situation above would be network B, as opposed to network A, the latter of which would have been selected by the original model proposed in [8].

4.1.1 Raw Reputation Score

In light of this observation, the first enhancement to the model proposed in [8] is for the reputation agent residing on the mobile node to compute a sample score which is not converted to a binary number. Instead, the raw reputation score is communicated to the network residing agent for aggregation. The purpose of this is that the raw score provides a more detailed representation of the service the mobile node received at the time the sample was calculated. In [8], the conversion of this score results in the loss of useful QoS information.

In addition to this, upon receipt of the sample score, the network residing reputation agent will then immediately compute the aggregate amount instead of computing the difference between the number of mobile nodes experiencing good service vs. those experiencing bad service. The reasoning behind this is that to be consistent it is preferable that each reported sample score communicated by a mobile node affects the

overall computed network reputation score in the same fashion. By first combining the scores in a given period, one mobile user's sample score in one interval could potentially have a different impact on the computed network score if the same score was leveraged in another interval. In this way, instead of combining the received scores within an interval, the network residing reputation agent in the proposed model queues sample scores and processes them sequentially in a first in first out (FIFO) manner. As a result, the sample score used as an input to the aggregate calculation can be seen in (4.1), with the aggregate computation already defined in (3.7),

$$r_{sample}(t) = \begin{cases} w^- Q_n(c_i), & Q_n(c_i) < Q_{th}(c_i) \\ w^+ Q_n(c_i), & Q_n(c_i) \geq Q_{th}(c_i) \end{cases}.$$

(4.1)

In this proposed model, a preferential weight can be given to positive and negative scores as well, in order to put more emphasis on one or the other. The score is considered to be positive when it is greater than or equal to a threshold, $Q_{th}(c_i)$, and negative when it is less than the same benchmark.

4.1.2 Applying GM(1,1) Prediction

Once the raw score has been run through the aggregation function in (3.7), instead of terminating the process and advertising the newly aggegated score to mobile nodes inquiring for access to the network, the proposed model takes one step further and utilizes the calculated aggregate score as an input into the GM(1,1) prediction model. More specifically, the network residing agent will store a window, n, of sequentially calculated aggregate scores and use them as values for the $X^{(0)}$ time series, represented by (3.8), which forms the basis of the GM(1,1) model,

This implies that on the subsequent receipt of a calculated sample score from a mobile node within the network, the network residing agent will recalculate the aggregate value and shift the persisted sequential aggregate scores to make room for the latest calculated value. As a result, in the previous cycle, the aggregate score which would have been assigned to $x^{(0)}(n)$ will now be shifted to $x^{(0)}(n-1)$, and the latest calculated value will assume the $x^{(0)}(n)$ position.

The result of this is that upon each receipt of a sample score, the network residing agent will as an end goal compute the predicted next aggregate score value and advertise this to the incoming mobile node users. The main benefit obtained from this is that mobile nodes wishing to obtain access to a given network will have advanced notice of network deteriorating below the service threshold of what is acceptable. The other perceived benefit is that spikes in service, either good or bad, should have a smoother effect on the predicted score since it will require multiple sequential values trending in a specific direction to significantly impact the overall predicted score. Finally, because this prediction calculation is done outside of the decision making process and within the network residing reputation agent only, the additional complexity introduced does not impact the delay incurred by the handover decision.

4.1.3 Polling Enhancement

The third major change made to the reputation based model proposed in [8] is an adjustment to the polling paradigm. In the original model, the mobile node devices compute a score only upon entry and exit of the network of service. Although mobile devices inherently are mobile and can travel into and out of a given coverage area relatively quickly, there is nothing preventing the mobile node from remaining in the

network for a reasonable amount of time, an example being a 3G cell phone user stopping for a coffee at a café and using a video chat service on his mobile phone. In this scenario, determining the score of the current network upon entry and exit leaves a large gap in the middle where the service may vary greatly and where the mobile user may desire to switch back to the 3G network if the option is available.

In order to mitigate this issue, and to provide more timely information about the service any given mobile user is experiencing within the network, the mobile nodes in the proposed model calculate a sample score of the given network periodically throughout the entire time they remain in the network, in addition to upon entry and exit of the network. The frequency of the sample reputation score calculation is determined based on two factors: the primary speeds a mobile user would maintain within an urban environment, and the average coverage zones of typical network technologies.

Since the calculation of the sample score is performed by the mobile node residing reputation agent outside of the decision making process, the increase in frequency of score calculation does not impact the delay incurred by the proposed model.

4.2 Proposed Vertical Handover Process

Similar to the methodology proposed in [8], the proposed model considers the scenario where the vertical handover either occurs by the express choice of the mobile node, i.e. the *passive handover* case, or simply out of need to maintain service continuity, i.e. the *forced handover* case. This overall process is defined in Figure 4.2 below.
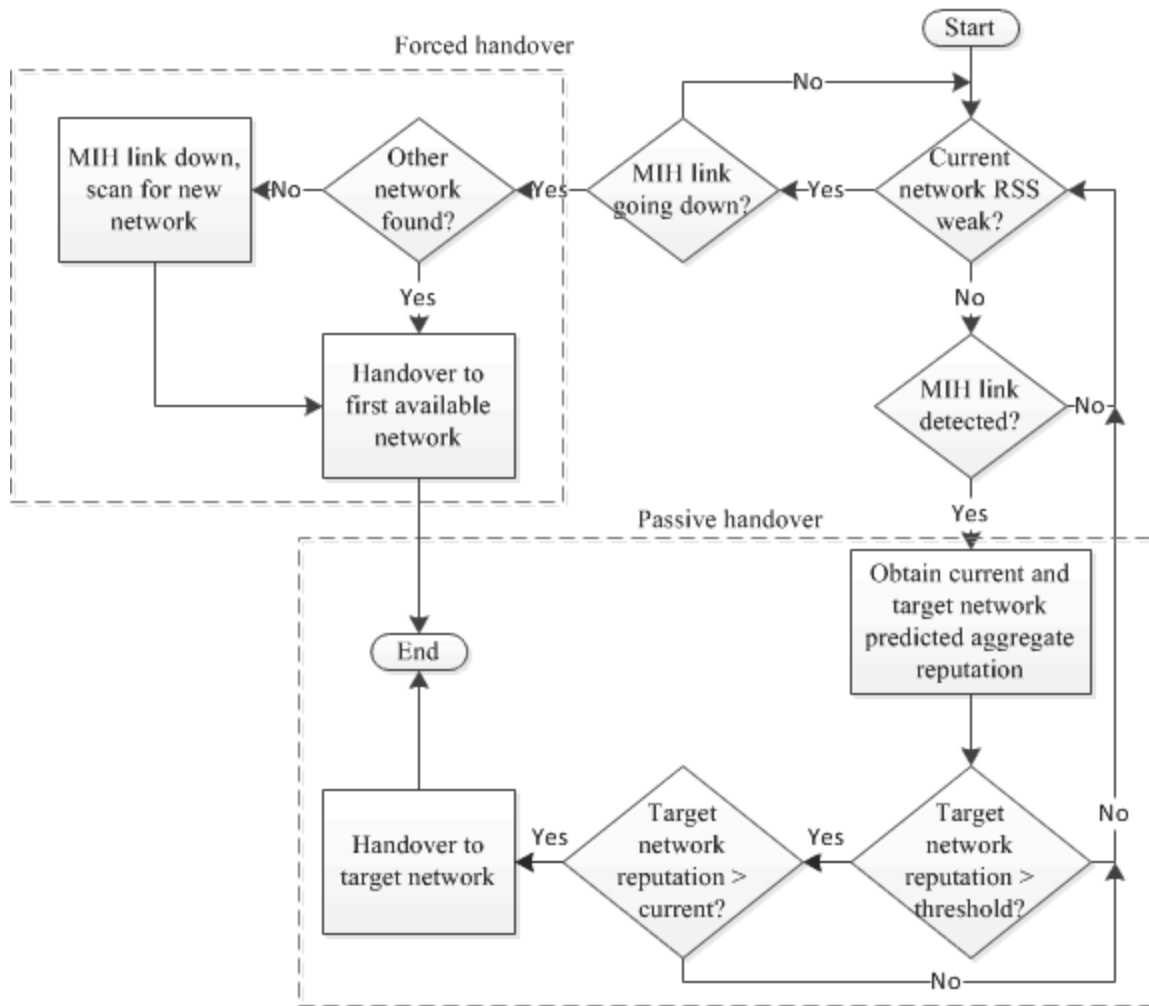
Figure 4.2: Proposed vertical handover decision algorithm process flow

4.2.1 Passive Handover

In the proposed model, the primary focus of investigation is within the passive handover scenario. In this case, the VHO decision is initiated based on the receipt of the 802.21 media independent handover event service *Link_Detected* event. The notification provided to the mobile device via this event is that a new network is within reach. As a result, the mobile node leverages its reputation agent residing on the network interface corresponding to the target network technology to reach out to the corresponding network residing agent to obtain the predicted score advertised by the target network.

Furthermore, the mobile node also uses the reputation agent residing on the current network interface to obtain the predicted score from the current serving network.

Once the predicted scores of the target and current network are obtained, the mobile device performs a quick comparison to see if the target network advertised score is both greater than the minimum threshold required for the service employed by the mobile device, and greater than the advertised score offered by the serving network. If both comparisons are favorable, the mobile device initiates the handover process using the available services from the MICS. Otherwise, if the predicted score offered by the target network is less than or equal to the threshold, or is less than or equal to the current network predicted score, the mobile device opts to not pursue the handover and remains in the current network. The reasoning behind not pursuing the handover when the predicted score of the target network is equal to the current network's predicted score is that there is no perceived benefit in making the handover. Since the scores are the same, and there is zero risk incurred by remaining in the current network, it is more sensible to refrain from making any changes.

4.2.2 Forced Handover

As in the passive handover case, the forced handover is initiated by a notification to the mobile device via the MIES that a link which the mobile device is currently using is either going down or has gone down, via the Link_Going_Down or Link_Down events.

Once this occurs, the mobile device immediately scans to determine if there is another network in the vicinity that can be leveraged. If one or more networks exist, the

mobile device will perform a handover to the first available network in order to ensure that some service is maintained. Otherwise, if there are no available networks to initiate a VHO to, the mobile device will continue to scan until a network is found. In the event that a mobile device makes a handover to a poorly performing network, it will resume the passive approach and wait until the MIH framework provides a notification that another network is within range before attempting to handover to a network which offers better services.

4.3 Summary

In synopsis, the key enhancement tabled by the proposed model is to treat the aggregate scores calculated by the network residing reputation agents as inputs to the GM(1,1) algorithm, in order to produce a predicted reputation score that can be advertised to enquiring mobile devices. Furthermore, through leveraging the raw reputation score, a more rich representation of the mobile node's QoS experience in the network is communicated to the network residing reputation agent, which in turn produces a predicted score that is more reflective of a QoS offered. In addition, through instructing the mobile devices to periodically calculate the reputation score of the network in addition to upon entry and exit of the network, the network residing reputation agent is receiving a more detailed and up-to-date picture of what level of service the network is offering. This in turn helps ensure that the scores advertised by the agent are accurate. Finally, the proposed model considers two scenarios; a passive handover case, when a new network is detected through the MIES and the mobile device investigates whether or not it is worthwhile to handover to the new network; and a forced handover

case, where the mobile node hands over to the first available network based on the notification from the MIES that the serving network is going down.

5 Chatpter 5: Simulation Configuration

5.1 NS2 Network Simulator

The performance of the proposed model has been carried out using the open source discrete event network simulator, NS-2, v2.29. This popular simulation tool has been available to the community since 1989 [31], and as a result there has been numerous contributions made to the code base of the simulator over time. The functionality of NS-2 is written in C++, while simulation configuration and execution is managed via an object oriented version of Tcl, called OTcl. As a result, there is typically an initial learning curve when getting started with the tool. In order to effectively use the simulator to verify the proposed vertical handover decision module, several contributed modules were leveraged, namely the NIST mobility module and the EURANE UMTS module.

5.1.1 NIST Mobility Module

The most important open source contributed code module leveraged in this thesis was the mobility module provided by the National Institute of Standards and Technology, NIST [32]. This module implements the 802.21 media independent handover framework, and provides support for this framework within the 802.11 (Wi-Fi), UMTS, 802.16 (WiMAX), 802.3 (Ethernet), and 802.15.1 (Bluetooth) modules. Specifically, the NIST mobility package defines the MIHF and the MIH user entities, provides the ability for neighbor discovery, and implements the MICS and MIES at the link layer and the MIH layer. It should be noted that the MIIS is not implemented in the NIST package, but is also not leveraged as part of the proposed decision module.

In order to implement a multi-interface mobile device in NS-2, the NIST module proposes an interesting scheme. Since the only way of indicating a mobile interface within NS-2 is through the instantiation of a node, NIST devised a virtual node which links together numerous nodes of different technology types in order to model a multi-interface mobile device. As a result, when the 'virtual node' mobile device approaches a new network, it determines if it has an available interface for the technology in order to facilitate a handover. If it does, and the decision module advises to initiate the handover, the virtual device connects to the new network and moves the flow of traffic from the old interface to the new one.

As a result of this provided module, the main development effort involved is the extension of the NS-2 simulator to define the network reputation scheme, namely mobile interface residing reputation agents and network residing reputation agents, and the functions that each of these agents carries out.

5.1.2 EURANE Module

The second open source contributed code that was leveraged and even extended by the NIST mobility module, was the Enhanced UMTS Radio Access Network Extension for NS-2 (EURANE) module that models UMTS [33]. Although this readily provides the ability to implement a 3G network within the network simulation, the drawback of the module is that the UMTS coverage area is perceived as being ubiquitous, so it is not possible to simulate the movement of the UMTS interface in respect to an access point for UMTS. As a result, the UMTS network serves as a starting point for the mobile nodes prior to navigating to the WiMAX and Wi-Fi networks.

5.1.3 MPEG Traffic Generation Module

Finally, in order to effectively model the scenario of streaming video or voice, a video traffic generator based on the Transform Expand Sample (TES) model of Moving Pictures Experts Group v4 (MPEG4) trace files was leveraged [34]. This traffic generator produces traffic that has the same first order and second order statistics as an original MPEG4 trace. This module was required due to the fact that out of the box, NS-2 comes equipped with only 4 kinds of traffic generators: exponentially distributed traffic, Pareto distributed traffic, constant bit rate traffic, and traffic based on a trace file.

5.2 Network Topology and Simulation Parameters

Once the open source contributed modules have been setup in the simulator, and the extensions carried out for the reputation model and GM(1,1) prediction model, the actual simulation itself must be configured in order to test the proposed algorithm. In this section the assumptions that have been taken are stated, important simulation parameters are discussed, and the network topology is reviewed.

5.2.1 Assumptions

In order to successfully implement the proposed model, several assumptions were made about the simulated environment and how best to realize the proposed model and original reputation scheme from [8]. These assumptions are stated here along with supporting discussion of how the assumptions were made. Final numerical values used to represent some of the information touched upon in the assumptions section is covered in the *Simulation Parameters* section.

5.2.1.1 Polling

In the proposed model, the polling that is carried out by each mobile node occurs immediately upon entry into the network, then periodically upon a defined interval, and finally upon exit from the network. On the other hand, in the original reputation model the polling carried out by the mobile nodes is undertaken only upon entry into and exit from the network. In order to ensure that the one-time sample upon entry into the network provides a good representation of the QoS experienced, the original reputation scheme is modeled as polling for the network score one second after entering the network. This delay is introduced to offset any noise associated with the initial joining of the node into the network.

5.2.1.2 Agent Messaging

Within the proposed and original models, the mobile node residing reputation agents and the network residing reputation agents are actively communicating with each other over the network. In this simulation, it is assumed that the reputation agents have their own management control channel over which communication is effected, and this channel is modeled as lossless, introducing no delay into the equation. The reasoning behind this is that the focus of the thesis is on the improvement of the selection of the appropriate network when presented with a handover possibility, not on optimizing the delay of the handover process. Furthermore, since the proposed enhancements are performed outside the decision algorithm, it is safe to assume that these enhancements will not modify the delay aspect of the original model.

5.2.1.3 Network Reputation Perceived Performance

During the simulation, there are instances typically at the beginning where a given network does not have any nodes residing in it. In order to define a baseline aggregate or predicted score of the network, it is assumed that in general the Wi-Fi network is preferred over both WiMAX and UMTS, and that WiMAX is in general preferred over UMTS. This assumption drives the default values for the network advertised score, and is made based on the assumption that the performance is typically highest and least costly in the Wi-Fi hotspot, whereas the lowest perceived quality of service is in the 3G network.

5.2.1.4 Score Comparison

Within the original reputation model, the scores which are calculated are converted to a binary system of [-1,1], and aggregated over an interval. In the proposed model, the scores are left as raw values and then aggregated upon receipt. As a result, in order to make a consistent comparison the original model's raw scores, pre-binary conversion, are obtain and run through the same aggregation calculation initially proposed. This enables are more straightforward comparison between the two models, as the scores then become measured on the same order. Furthermore, it is assumed that the scores are processed sequentially and that there is no defined interval over which the network residing agent needs to collect and combine scores. Each received score is processed upon receipt – in the original model, this is processed via aggregation, and in the proposed model it is processed via aggregation and then prediction.

5.2.2 Simulation Parameters

5.2.2.1 Performance Metric Weights

In order to calculate the sample scores in the proposed algorithm and the original reputation model, weights need to be defined for the performance metrics used in the calculation of the score, as seen in (3.5).

In order to facilitate this, the calculations performed in [35] are leveraged, where the authors define the AHP matrices for 4 classes of services, as seen in the table 5.1.

Table 5.1: AHP matrix for conversational, streaming, interactive, and background CoS

Conversational	BER	Delay	Jitter	Bandwidth
BER	1	1/9	1/9	1
Delay	9	1	1	9
Jitter	9	1	1	9
Bandwidth	1	1/9	1/9	1

Streaming	BER	Delay	Jitter	Bandwidth
BER	1	1/5	1/9	1/9
Delay	5	1	1/5	1/5
Jitter	9	5	1	1
Bandwidth	9	5	1	1

Interactive	BER	Delay	Jitter	Bandwidth
BER	1	5	9	5
Delay	1/5	1	5	1
Jitter	1/9	1/5	1	1/5
Bandwidth	1/5	1	5	1

Background	BER	Delay	Jitter	Bandwidth
BER	1	9	9	9
Delay	1/9	1	1	1/5
Jitter	1/9	1	1	1/5
Bandwidth	1/5	5	5	1

As a result of these matrices, the actual weights used in the score calculations are defined in [35] and seen in table 5.2, with an associated consistency ration (CR).

Table 5.2: Weights calculated in [35] for each performance metric, per CoS

Traffic Class	BER	Delay	Jitter	Bandwidth	CR
Conversational	0.04998	0.45002	0.45002	0.04998	0.000
Streaming	0.03737	0.11380	0.42441	0.42441	0.049
Interactive	0.63593	0.16051	0.04304	0.16051	0.049
Background	0.66932	0.05546	00546	0.21976	0.049

5.2.2.2   Metric Normalization and Threshold Calculation

Based on the weights defined in [35], the score threshold against which aggregate/predicted scores are compared against can be calculated, given that the specific performance metric normalization and threshold values are defined. Since the primary traffic under investigation in this thesis is related to videoconferencing or streaming music/video data, these metrics are defined for the conversational CoS.

Beginning with the threshold values, in order to assign proper values for each performance metric, the applicable minimal value required to maintain an acceptable level of QoS is utilized. Based on typical implementations [36], [37], these values are determined as follows:

- **BER**:   60 packets/s

- **Delay**:   150ms

- **Jitter**:   10ms

- **Bandwidth**:   2Mbps

Secondly, the normalization factors are determined in two ways. For the simulations against a single mobile node, the normalization metrics are determined based on average values known for the given CoS in order to make some distinction between

the normalization and threshold values. This results in the following normalization values, and a calculated threshold of 0.026.

- **BER**: 1 packet/s

- **Delay**: 1ms

- **Jitter**: 0.1ms

- **Bandwidth**: 5.5Mbps

For the simulations involving 60 nodes, the normalization values are determined empirically through gradually overloading the Wi-Fi and WiMAX networks separately. In this fashion, the highly congested network performance metrics are observed and the normalization weights are defined as:

- **BER**: 0.01 packets/s

- **Delay**: 10ms

- **Jitter**: 1ms

- **Bandwidth**: 3Mbps

As a result of the above parameters, the threshold score calculated in the simulation for the conversational class of service is defined to be 0.075.

5.2.2.3 Network Reputation Score Defaults

As indicated in the assumptions, the three networks modeled in the simulation are, in the order of preference, Wi-Fi, WiMAX, and UMTS. In order to distinguish this distinction in perceived service between the three network types, each are assigned a default network score as follows:

- **Wi-Fi**: 2

- **WiMAX**: 1

- **UMTS**: 0.08

Since UMTS is perceived as the lowest favorable network to join, the default network score is assigned a value slightly above the threshold value of 0.075, i.e. 0.08. The other two networks are then arbitrarily set to 1 and 2 to facilitate the preference.

5.2.2.4 Score Aggregation Weights

In addition to the AHP calculated parameters utilized in the sample score calculation, there are two additional weights that are leveraged by the network residing reputation agent when performing the aggregation, as seen in (3.6) through w^+ and w^-, and in (3.7) through γ.

In order to assign higher importance to the below-threshold observed scores and to the most recently obtained scores, the weights are assigned as follows:

- $w^+ = 0.4$

- $w^- = 0.6$

- $\gamma = 0.6$

5.2.2.5 Polling and Mobile Node Speed

The environment that the simulation is attempting to model is a dense urban environment. As a result of this, it is expected that the mobile node can typically travel at three common speeds: walking speed, non-highway city speed (e.g. riding a bus), and highway speed (e.g. riding in a car). This equates to the following speed levels:

- Walking: 5km/h, or 1.38 m/s

- Non-highway (Bus): 40km/h, or 11.11 m/s

- Highway (Car): 75km/h, or 19.4 m/s

Given that the minimum range of coverage between the three networks is the Wi-Fi technology, and assuming that this coverage could have a range of 250m [4], a mobile node travelling at the highest speed could sample 2.57 times within the Wi-Fi network if the polling interval was allocated to a period of 5s. As a result, this frequency is used in the proposed model.

5.2.3 Network Topology

For the single node simulation, the topology which is used is defined as seen in Figure 5.1, where a single mobile node interacts with a UMTS, WiMAX, and Wi-Fi network. The Wi-Fi network is partially overlaid into the WiMAX network, and the UMTS coverage is ubiquitous. One traffic source is attached to router one and sending the MPEG traffic to the mobile node.
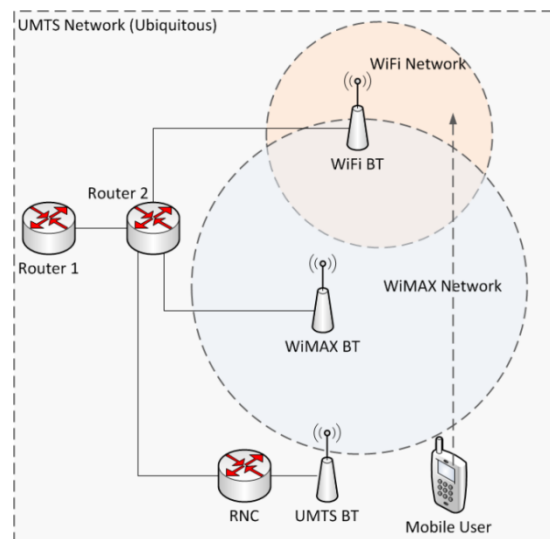
Figure 5.1: Network topology modeled for testing VHO with 1 mobile node

In order to measure the performance of the proposed model vs. the original model in a multiple node configuration, the following network topology is leveraged.
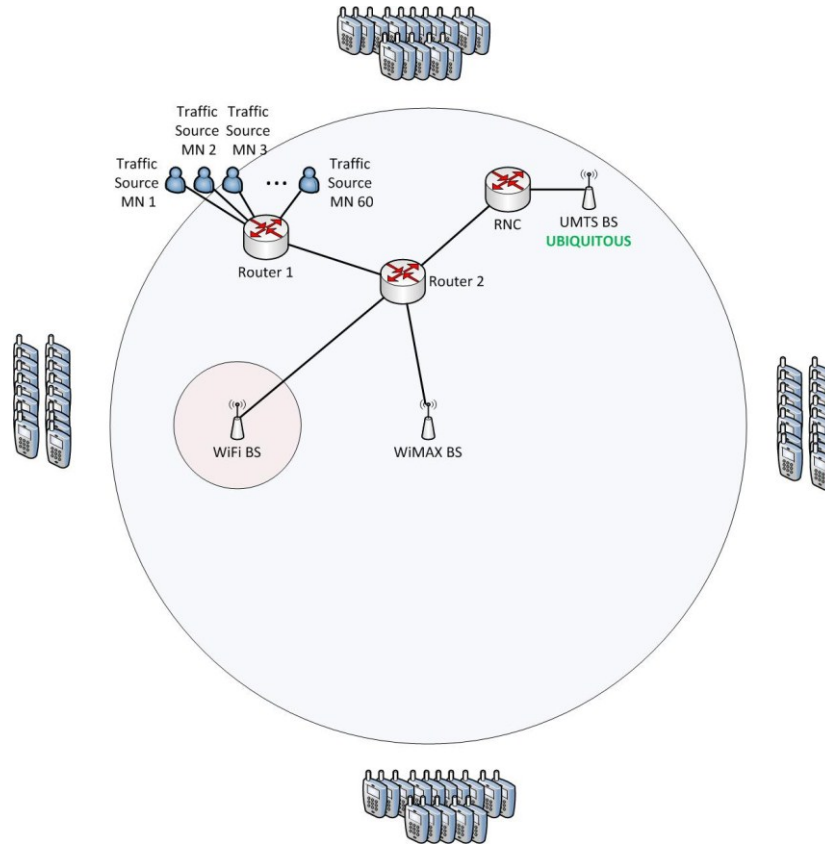
Figure 5.2: Network topology for testing VHOs with multiple mobile nodes

In this topology, 60 multi-interface mobile devices begin within UMTS, whose coverage is ubiquitous. The WiMAX base station is located at the (1000, 1000) coordinate, and the Wi-Fi base station is located within the WiMAX coverage at the coordinate (680, 1000). Sixty traffic sources are deployed in the network and each is paired with a mobile node in order to deliver it the video content; these traffic sources are attached to the core network router 1, which routes through the gateway router 2 to the appropriate base station based on the destination of the mobile node.

5.3 Summary

In this chapter, the assumptions pertaining to the polling of the mobile node residing reputation agents, the reputation agent messaging, the network technology perceived preference, and the score comparison is discussed. Furthermore, the simulation configuration parameters that help drive the experiment are also stated, namely the actual AHP calculated weights for the sample score calculation, the weights used for threshold setting and performance metric normalization, the network reputation score defaults, the aggregation weights, and the polling interval. Given this set of information, one is prepared to review the simulation results.

6   Chapter 6: Performance Evaluation

6.1   Single Node Observations

6.1.1 Gradually Varying Traffic

The first experiment involving a single node considers the gradual increase and decrease of traffic as a node traverses from one network into another. In this scenario, the traffic applied to the Wi-Fi network is controlled in order to produce gradually varying scores. The mobile node is iteratively made to traverse from the WiMAX to the Wi-Fi zone at different times throughout the varying load curve, as seen in Figure 6.1.

Figure 6.1: Comparison of sampled reputations score over time between proposed prediction model (red) and existing aggregate model (blue), with gradually varying traffic load

From the Figure, it can be seen that as the network becomes congested the mobile node using the predicted score at time (t) has advanced warning that the QoS of the network is deteriorating. This can be seen at time t = 35s and t = 40s, where the

aggregate scores are above the threshold and perceived as good, while the predicted scores are below the threshold and perceived as bad. Since the aggregate scores following the above time instances do fall beneath the threshold, the predicted scores allow the mobile node to make the VHO decision sooner and maintain an increased overall experience.

6.1.2 Traffic Spike

The second experiment carried out against a single node considers a spike occurring within the traffic flow. In this scenario, the traffic applied to the Wi-Fi network is controlled to produce gradually varying scores with a sudden surge introduced temporarily and then removed. Similarly to the first experiment, the mobile node is iteratively made to traverse from the WiMAX to the Wi-Fi zone at different times throughout the varying load curve, as seen in Figure 6.2.
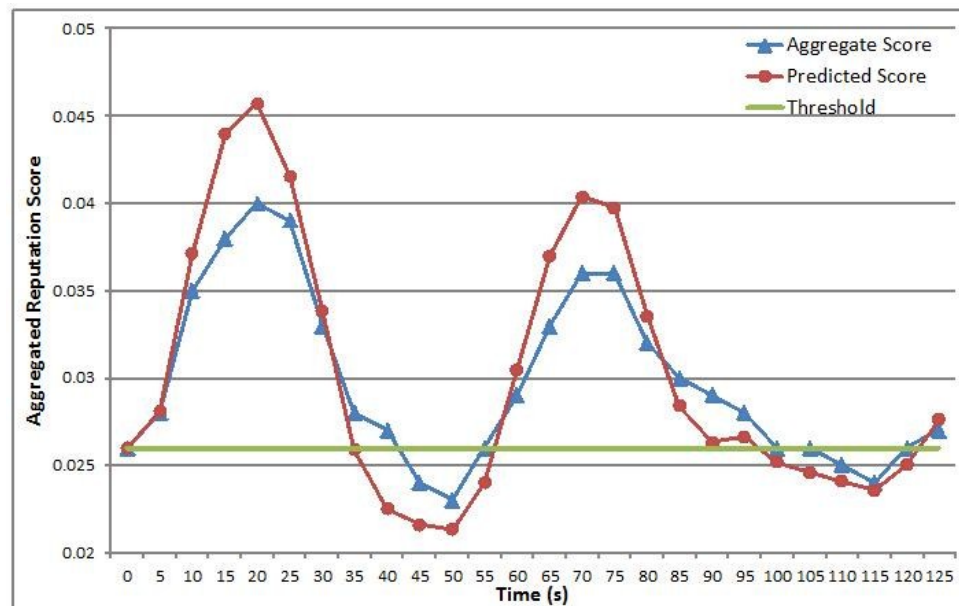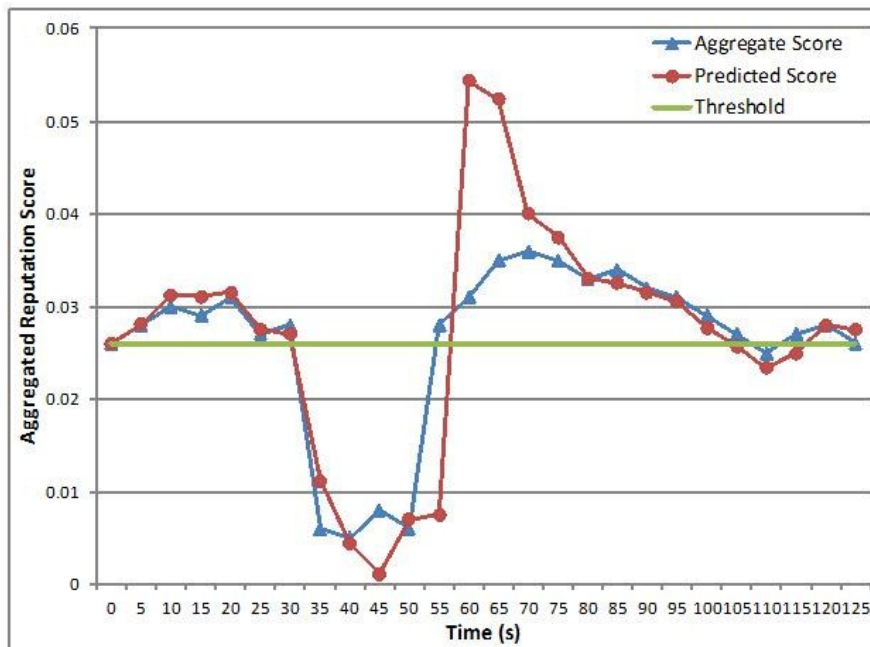

Figure 6.2: Comparison of sampled reputations score over time between proposed prediction model (red) and existing aggregate model (blue), with traffic load spike

As is evident from the figure, the spike in traffic corresponds to the sudden drop in the aggregate score at t = 35s. Furthermore, it is clear that the predicted score does not provide any additional benefits in the case of a sudden degradation of the network. This makes sense as there are no prior values that can give advanced warning of the sudden change in traffic load. The advantage can be seen when the QoS is re-established at time t = 55s. Instead of deciding to join on the first instance of a perceived good score, as in the case of the aggregate score, the predicted score requires the occurrence of at least 2 'good' scores. This enables the mobile node to make a more confident VHO decision and also help reduce the number of unnecessary handovers overall in the network.

6.2 Multi-Node Observation: Overload WiMAX and Wi-Fi

In order to assess the improvements that the proposed model yields, it is compared against the original reputation scheme from [8] against two types of scenarios: one pure overload scenario, and one with a mix of overload and decrease load. As mentioned in section 5.2.1.4, Assumptions > Score Comparison, the pre-binary conversion score is obtained from the original reputation scheme and run through the aggregation process in order to compare scores which are on the same order.

6.2.1 Scenario Description

The purpose of the first scenario is to sequentially overload the WiMAX and then the Wi-Fi networks. Based on the decision algorithm defined in the proposed model, the expected behavior is that there will be fewer handovers in the proposed model as compared to the original model, since the proposed model should be able to sense earlier

that the network is becoming congested. The flow of events in this scenario, at a speed of 40km/h, is described as follows, referring to Figure 6.3:

1. Four groups of 15 mobile nodes reside outside the four quadrants of a WiMAX base station, which is located at (1000, 1000).

2. At 10s, the first group located closest to the Wi-Fi base station, begins moving towards the Wi-Fi BS, with one node beginning to move every second.

3. At 15s, the group below the WiMAX BS begins moving midway into the WiMAX region, with one node moving every second.

4. The same occurs for the groups to the right of and above the WiMAX BS at 20s and 25s respectively. At this point the WiMAX network is becoming overloaded.

5. At 60s, the transition to the Wi-Fi network begins, with a mobile node from the group to the right of the WiMAX BS beginning to move toward the Wi-Fi BS every 2s

6. At 65s, again every 2 seconds, a node from the groups above and to the south of the WiMAX BS begin moving toward the Wi-Fi BS as well, overloading the Wi-Fi coverage zone.

7. The simulation ends after 250s.

This scenario is repeated for the mobile speeds of 75km/h and 5km/h, and in order to ensure that the scenario is the same, the timings are adjusted to compensate for the speed change. For instance, in the case of 5km/h, the overall simulation time is 1000s instead of 250s.
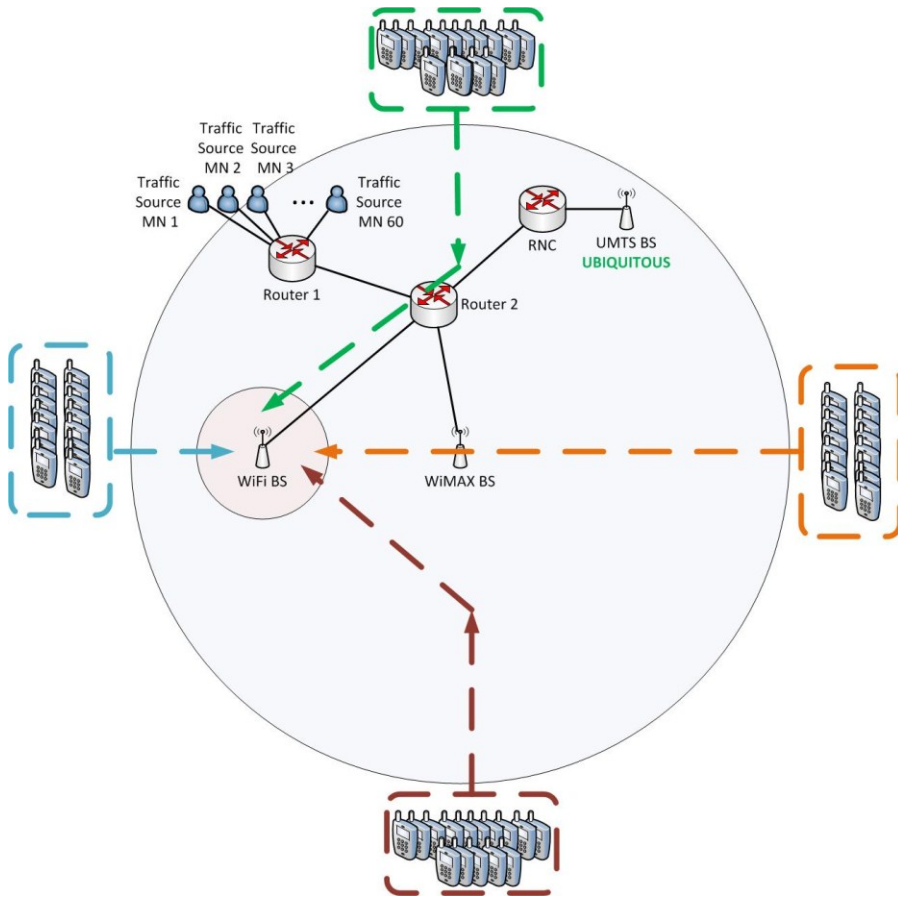
Figure 6.3: Topology and mobile node path for WiMAX and WiFi traffic overload scenario

6.2.2 WiMAX Results

In this section, the results of the three different speeds are reviewed as they pertain to the WiMAX coverage area. In general, the comparison between the total number of handovers realized upon completion of the simulation within the proposed model and original model align to the expected behavior, as can be seen in Figure 6.4.
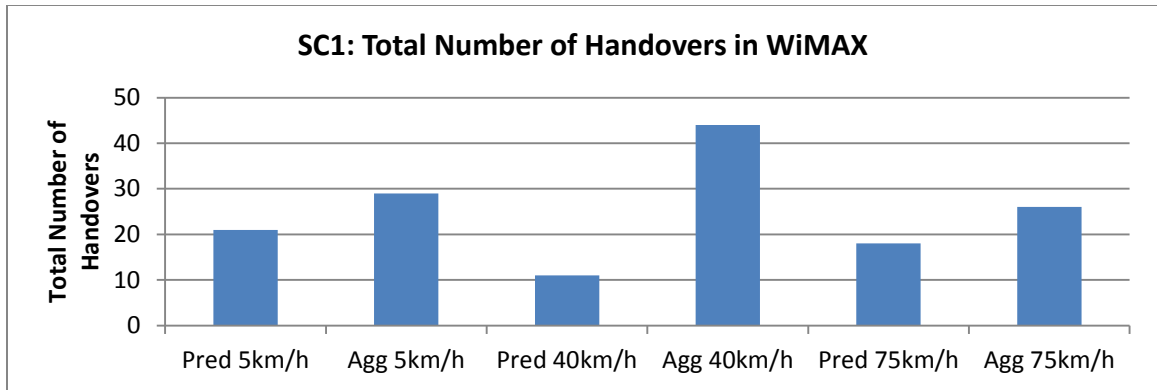
Figure 6.4: Scenario 1 comparison of the total number of VHOs between the proposed prediction model and the existing aggregate model at mobile node speeds of 5km/h, 40km/h, and 75km/h in WiMAX

6.2.2.1   Mobile Speed 5 km/h

Based on the description of the scenario, the four different groups of mobile nodes are beginning to move toward the WiMAX network every 5 seconds at a rate of 1 mobile node per second.   Since there are more than 5 nodes per group, there are concurrent arrivals at the WiMAX coverage area, specifically every 5s.   As a result, the score advertised by the network reputation agent is gradually decreasing due to the increasingly congested network, as seen in Figure 6.6.   Given this, the rate of handovers between the two protocols can be observed in Figure 6.5.

Figure 6.5: Scenario 1 comparison of the cumulative # of handovers over time between the proposed prediction model (blue) and the existing aggregate model (red) at 5km/h in the WiMAX network

Based on this depiction, it is evident that the total number of permitted handovers up until 56s is the same between the prediction and aggregate methods, but at 60s the total number of handovers permitted by the aggregate increases by 14, while in the prediction model it only increases by 4. This is due to the advanced notice provided by the proposed model that the threshold will soon be breached. As a result, it can be noted that the sample scores post 56s are significantly better than the aggregate method, as seen in Figure 6.7; as expected, this behavior is also reflected in the advertised network scores (Figure 6.6).

Figure 6.6: Scenario 1 comparison of the reputation scores perceived by the mobile nodes upon entering the WiMAX network at 5km/h, between the prediction model (blue) and aggregate model (red)

Figure 6.7: Scenario 1 comparison of the reputation scores sampled by the mobile nodes while within the WiMAX network at 5km/h, between the prediction model (blue) and aggregate model (red)

6.2.2.2 Mobile Speed 40km/h:

At 40km/h, one can observe that after 20s there are significantly more handovers that occur in the aggregate model as compared to the prediction model. As a result, the sample scores calculated by the mobile nodes in the network reach a lower score faster in the aggregate model than in the prediction, as seen in Figure 6.9 below. At time 22s the aggregate model has already reached a fairly low score of approximately 0.58, whereas the prediction model only gets close to that level around 32s. This is due to the predictive nature of the proposed model, as seen when the speed was 5km/h.
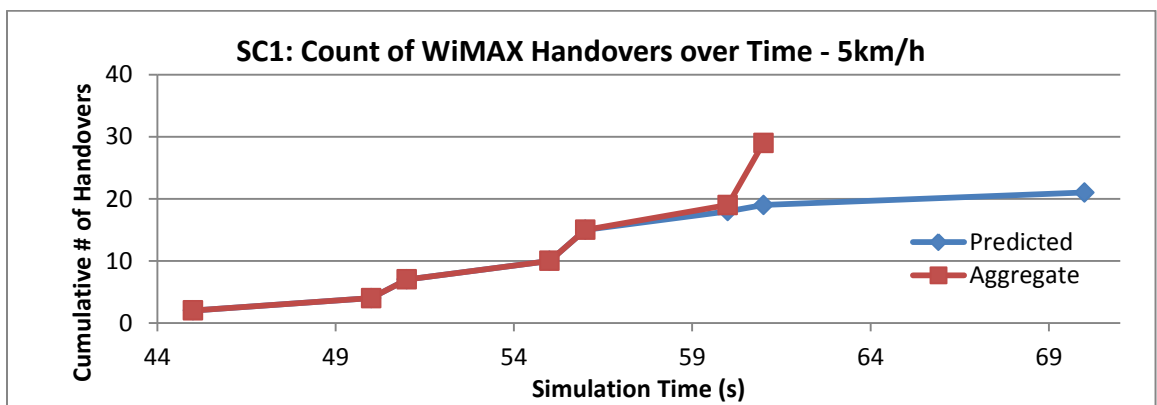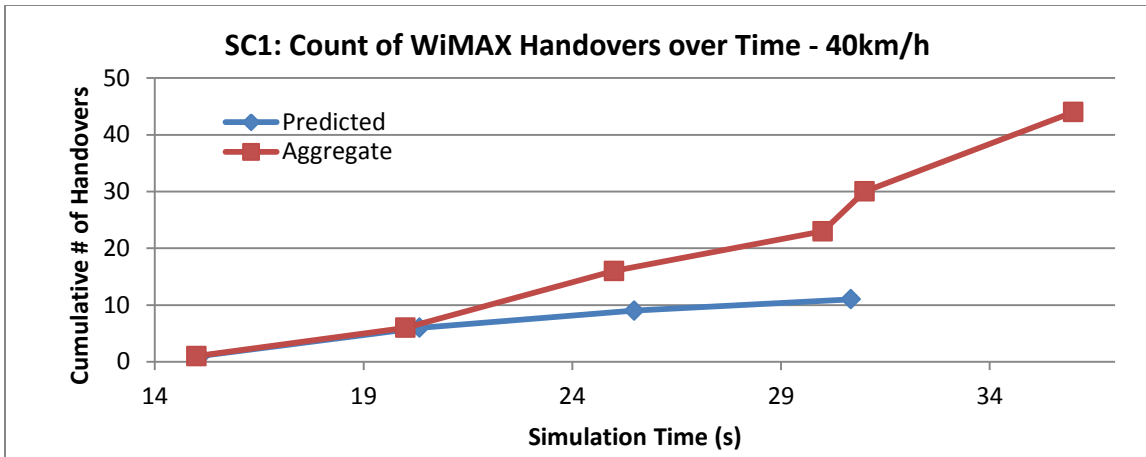
Figure 6.8: Scenario 1 comparison of the cumulative # of handovers over time between the proposed prediction model (blue) and the existing aggregate model (red) at 40km/h in the WiMAX network
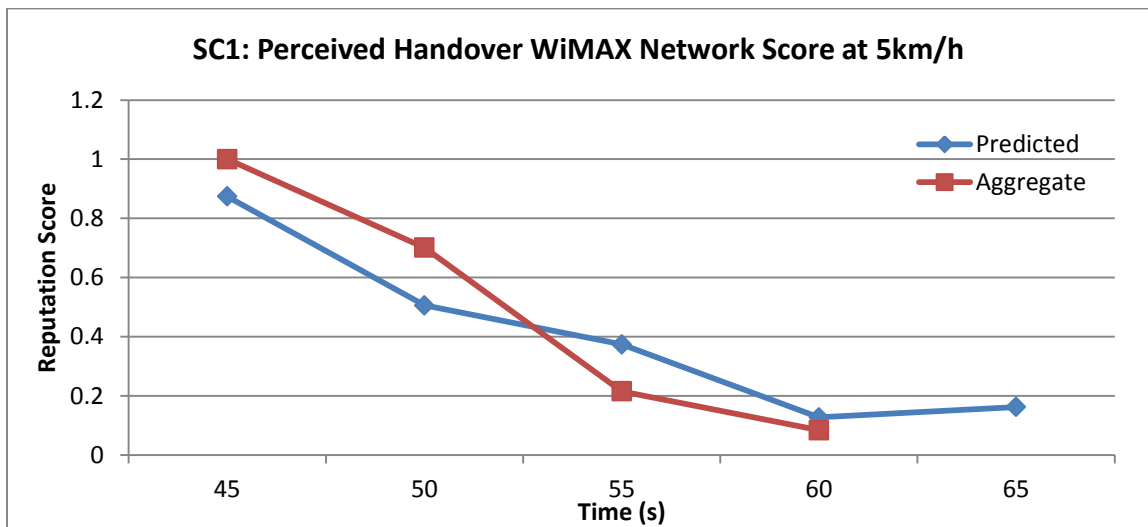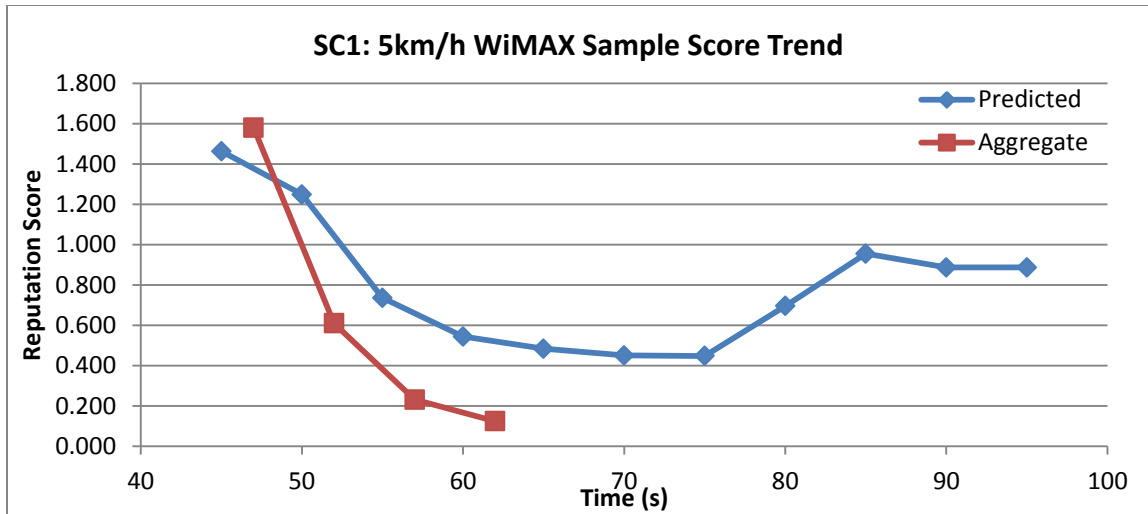

Figure 6.9: Scenario 1 comparison of the reputation scores sampled by the mobile nodes while within the WiMAX network at 40km/h, between the prediction model (blue) and aggregate model (red)

From the sample scores in Figure 6.9, the minor oscillation seen in the aggregate model at 32s and at 55s in the prediction method is due to one mobile node receiving a significantly higher QoS than the other nodes, which could be due to resource contention in the network. The minor peaks in the aggregate model at 34s, 35s, and 36s, are due to the fact that some of the nodes from the blue group are beginning to join the Wi-Fi

network, but this initial offload of traffic is quickly counterbalanced by the new nodes arriving into the network from the other three groups.

In reviewing the handover scores that are perceived by the mobile nodes that join the WiMAX network, as depicted in Figure 6.10, there appears to be a spike at 20s for the aggregate method. This is explained as follows: the default score defined for the network residing agent is 1. At 15s, 1 node joins the network (via both methods). In the aggregate method, the score is only calculated once, and this is done *1s after* joining the network. In the prediction method, the score is calculated immediately, and then re-calculated every 5s. In this way, at 15s, both the aggregate and the prediction algorithms perceive a network score of 1. The difference is that at 20s, when 5 other nodes join the network (in both methods) the prediction model has each node immediately calculate the score and send it to the network node for aggregation / prediction. Due to the congestion of 5 nodes immediately joining the network and the score being calculated within each other's presence felt, the average network score result is less than 1. In the aggregate model, the score is only calculated by each node 1 second after 20s, which is why the score seen at 20s is only the score calculated by the 1 node which joined at 15s.
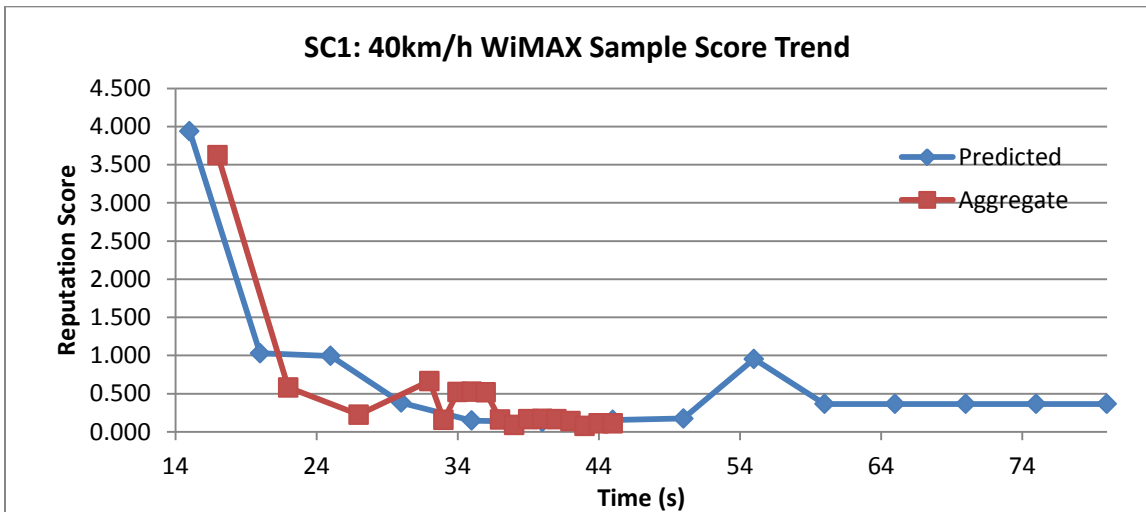

Figure 6.10: Scenario 1 comparison of the reputation scores perceived by the mobile nodes upon entering the WiMAX network at 40km/h, between the prediction model (blue) and aggregate model (red)

6.2.2.3 Mobile Speed 75km/h:

At the top mobile node speed, the pattern follows suit with what was seen in the 40km/h and 5km/h scenarios. At 25s we see a divergence in the amount of handovers that are permitted between the aggregate and prediction model, with the aggregate allowing roughly ten more cumulatively over the remainder of the simulation, as seen in Figure 6.11. Consequently, this results in the mobile nodes calculating higher sample scores within the prediction model after 26s, as seen in Figure 6.12.

Figure 6.11: Scenario 1 comparison of the cumulative # of handovers over time between the proposed prediction model (blue) and the existing aggregate model (red) at 75km/h in the WiMAX network

Figure 6.12: Scenario 1 comparison of the reputation scores sampled by the mobile nodes while within the WiMAX network at 75km/h, between the prediction model (blue) and aggregate model (red)

Similarly to the scenario at 40km/h, there are some minor peaks in the aggregate model around 24s and 30s – this is due to the fact that some of the nodes from the blue group have begun joining the Wi-Fi coverage zone beginning around 22s. Due to some nodes leaving in conjunction with some nodes arriving, the peaks appear intermittently but taper off as the last node enters the Wi-Fi network around 36s.

Through the perceived network advertised scores seen in Figure 6.13, both the aggregate and prediction model follow a similar curve, due to the degradation of the network. It can be noted that at this speed, there is no peak seen after the initial value in the aggregate mode, as was seen at 40km/h. This is due to the fact that 3 nodes initially join at the beginning instead of 1, providing a more congested score calculated after the first second.
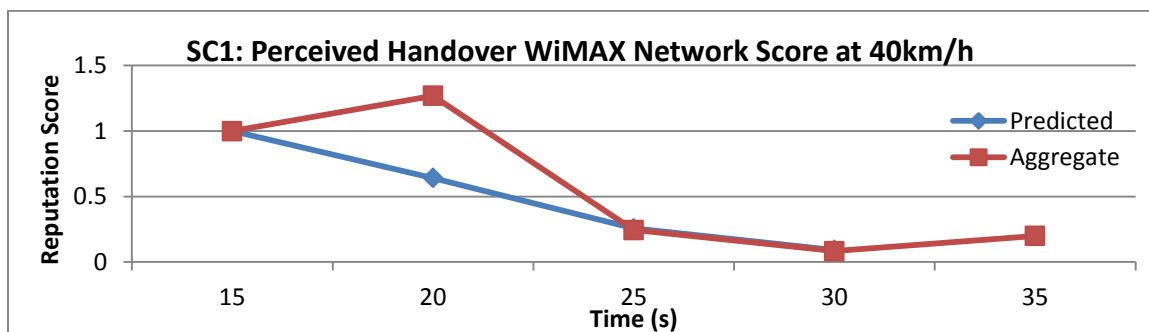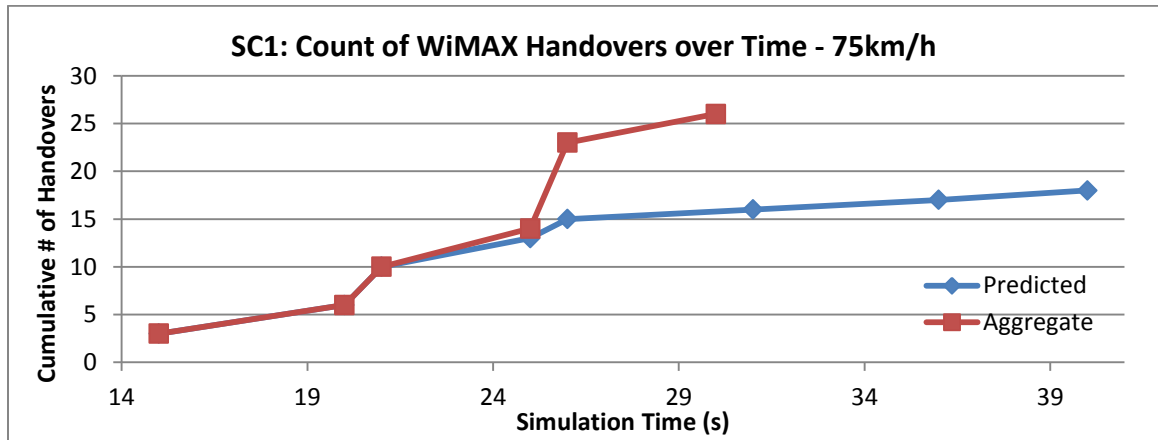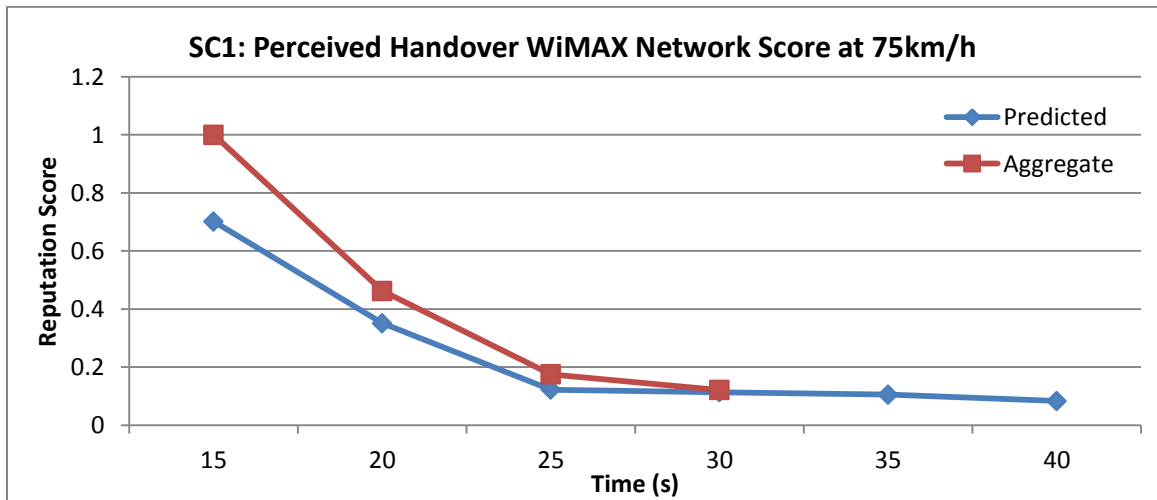

Figure 6.13: Scenario 1 comparison of the reputation scores perceived by the mobile nodes upon entering the WiMAX network at 75km/h, between the prediction model (blue) and aggregate model (red)

6.2.3 Wi-Fi Results

Subsequently to overloading the WiMAX network, the mobile nodes converge on the Wi-Fi network, causing congestion. In general, since the network is being gradually congested, the expectation is that there would be less handovers in the prediction model. For the most part this is true, as can be seen in Figure 6.14; the exception at 40km/h is discussed in addition to the other results in this section.
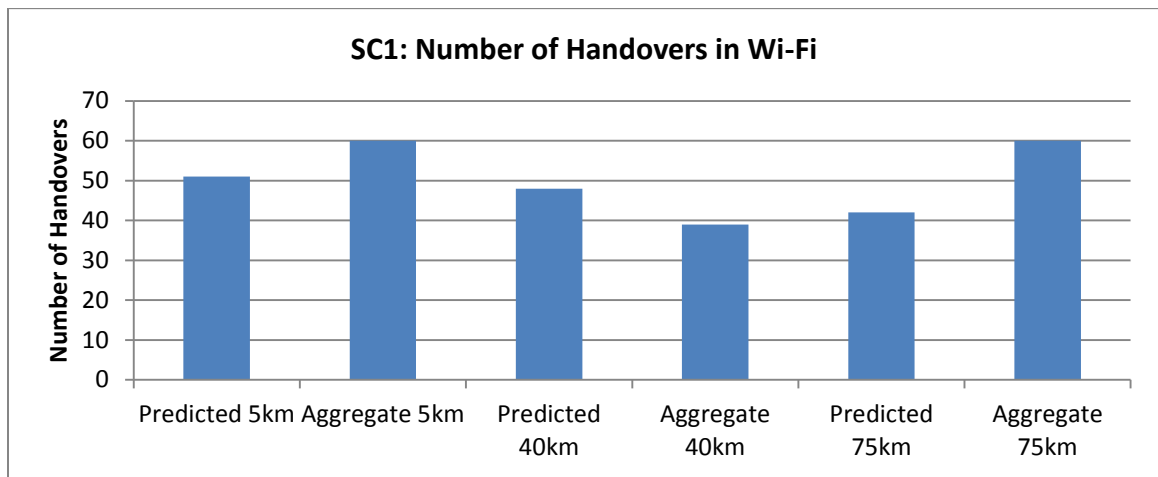
Figure 6.14: Scenario 1 comparison of the total number of VHOs between the proposed prediction model and the existing aggregate model at mobile node speeds of 5km/h, 40km/h, and 75km/h in Wi-Fi

6.2.3.1 Mobile Speed 5km/h:

At this speed, because the mobile nodes are moving slowly, they arrive at three different time periods. Due to the gradually incremental load, the total number of handovers over time is less in the prediction model, as can be seen in Figures 6.15, 6.16, and 6.17, where the aggregate begins to diverge from the prediction around 440s.

Figure 6.15: Scenario 1 comparison of the cumulative # of handovers over time between the proposed prediction model (blue) and the existing aggregate model (red) at 5km/h in the Wi-Fi network; first period of arrival of mobile nodes
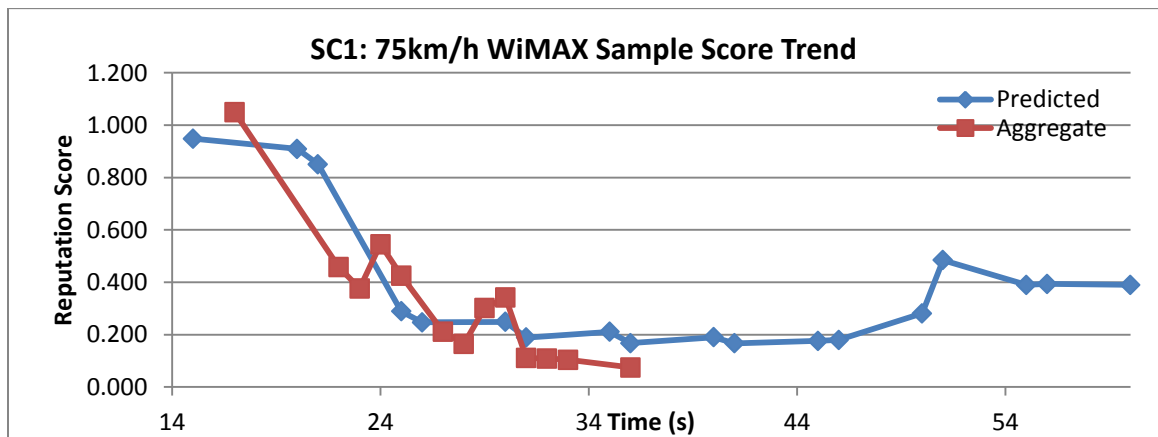

Figure 6.16: Scenario 1 comparison of the cumulative # of handovers over time between the proposed prediction model (blue) and the existing aggregate model (red) at 5km/h in the Wi-Fi network; second period of arrival of mobile nodes
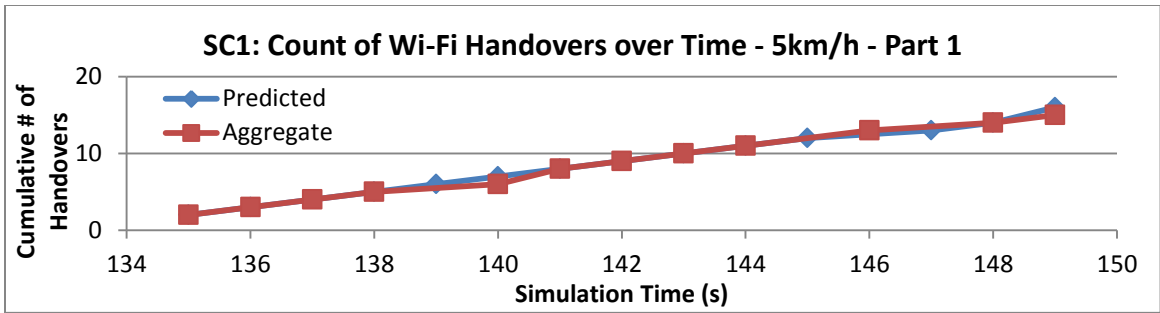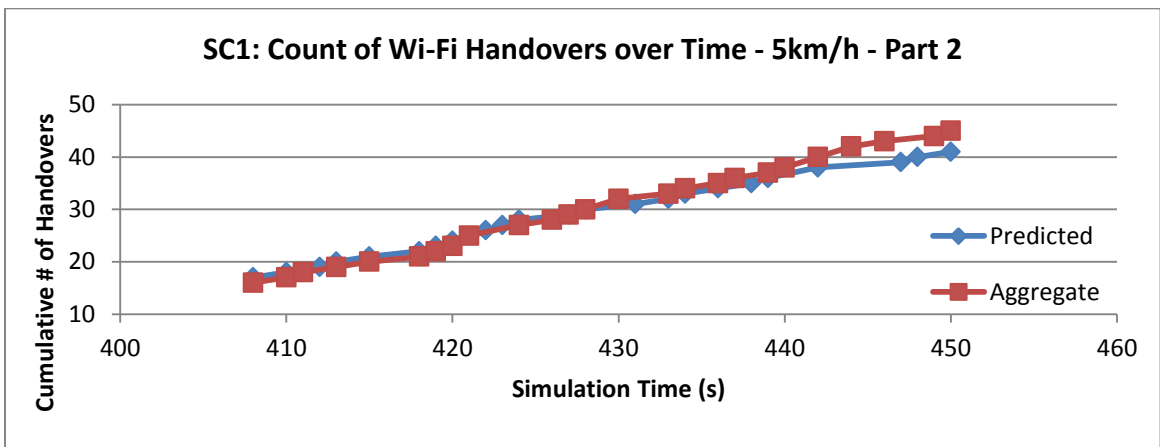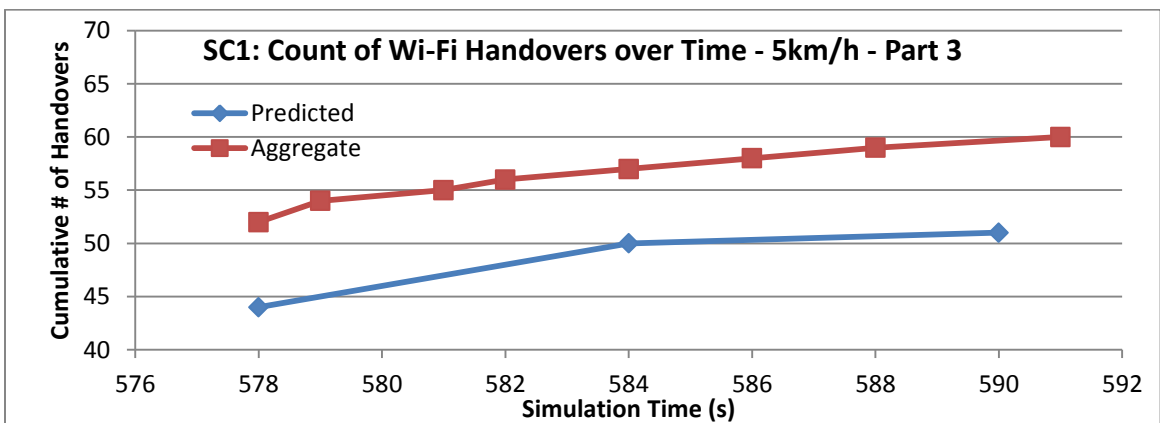

Figure 6.17: Scenario 1 comparison of the cumulative # of handovers over time between the proposed prediction model (blue) and the existing aggregate model (red) at 5km/h in the Wi-Fi network; third period of arrival of mobile nodes

This behavior can be seen mirrored in the sample scores polled by the mobile nodes in the network, in Figures 6.18, 6.19, and 6.20, where the majority of the scores in the prediction model after 440s are below the aggregate model up until the last node from the second wave joins at 450s. after this point, and until the third wave of nodes begin joining, the predicted scores stabilize around 2.5 since there are no additional nodes joining and causing the predicted score to increase or decrease (note that there are no aggregate sample scores after 450 because no other nodes join). The numerous spikes that are observed in the sample score trends are due to the fact that some of the mobile nodes, upon entry into the network, manage to obtain significantly higher QoS from the network. The longer they remain in the network, the more their grasp on the network resources decreases and ends up aligning to the average experienced by the other nodes in the network, as seen by the decreasing peaks in Figure 6.19. This is attributed to the fact that there is no resource reservation control in the network, so contention between the nodes for resources is constantly occurring.

Figure 6.18: Scenario 1 comparison of the reputation scores sampled by the mobile nodes while within the Wi-Fi network at 5km/h, between the prediction model (blue) and aggregate model (red); first period of arrival of mobile nodes

Figure 6.19: Scenario 1 comparison of the reputation scores sampled by the mobile nodes while within the Wi-Fi network at 5km/h, between the prediction model (blue) and aggregate model (red); second period of arrival of mobile nodes

Figure 6.20: Scenario 1 comparison of the reputation scores sampled by the mobile nodes while within the Wi-Fi network at 5km/h, between the prediction model (blue) and aggregate model (red); third period of arrival of mobile nodes

Another contributing factor to more nodes joining the Wi-Fi network in the aggregate model is that in the prediction model, mobile nodes are constantly polling for the score and as a result, the network score is always being recalculated. In the aggregate

model, the score is only calculated upon entering the network and when leaving due to forced handover related reasons. As a result, when the mobile nodes are experiencing congestion in the Wi-Fi network between 440s and 450s, the WiMAX network perceived by the prediction model is experiencing generally better service, as seen in Figure 6.21.
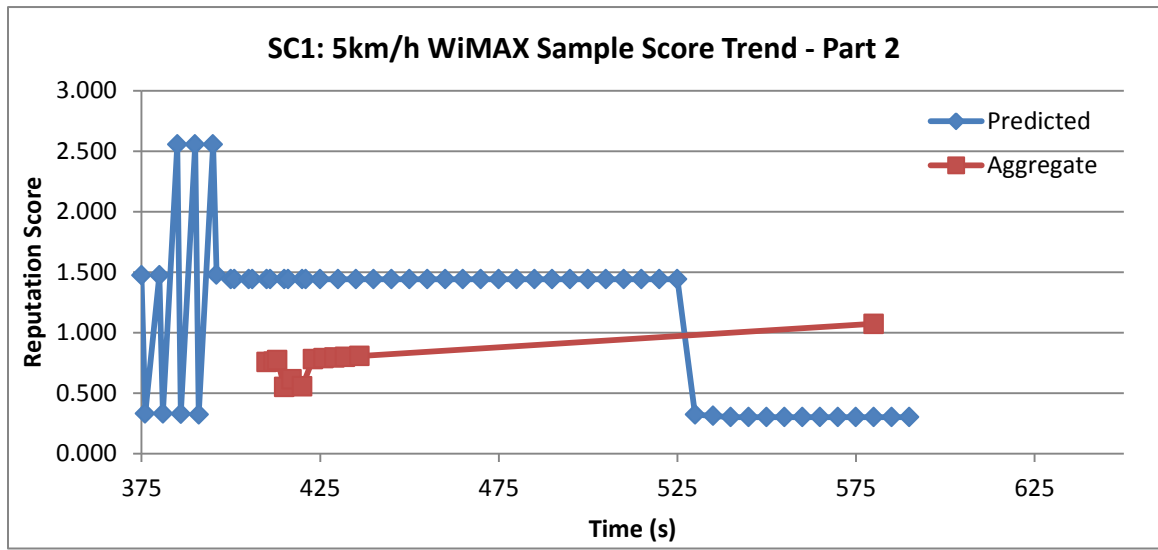
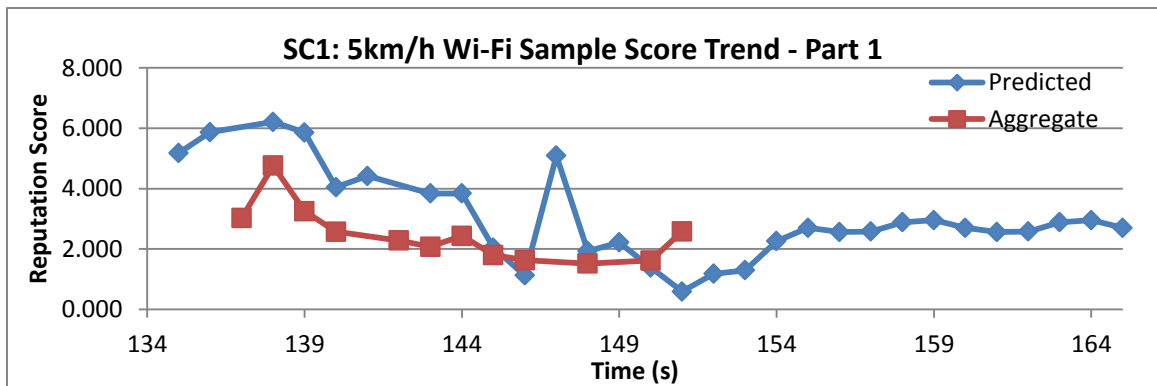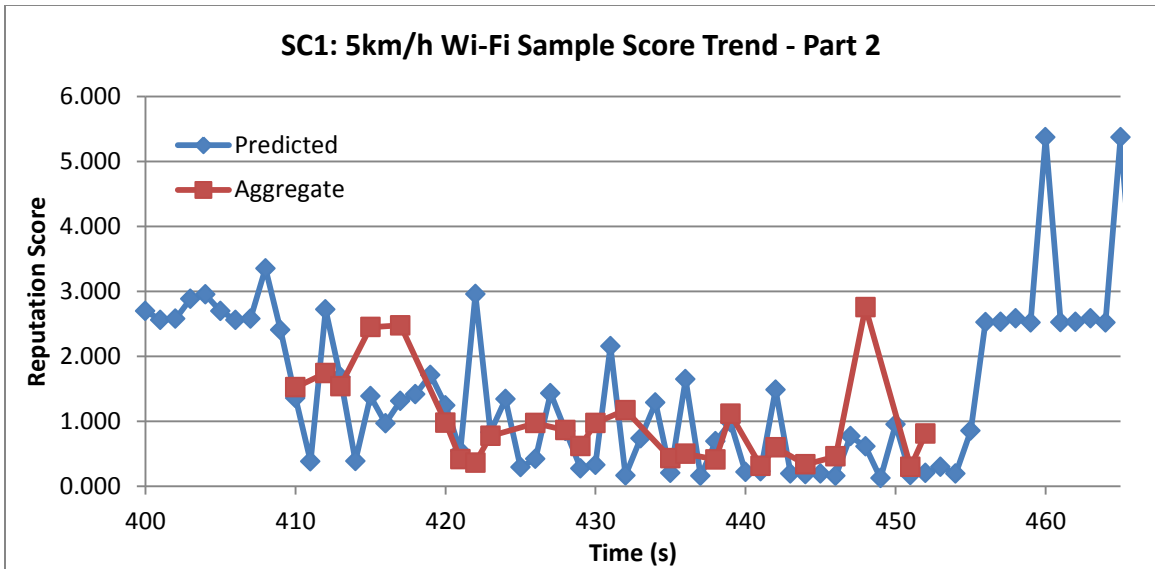**Figure 6.21: Scenario 1 comparison of the reputation scores sampled by the mobile nodes while within the WiMAX network at 5km/h, between the two models; during this period the second and third wave of mobile nodes are presented with the option to handover to Wi-Fi**

During the period between the second and third wave, the oscillations observed are due to one of the mobile nodes having obtained a higher QoS in the network. Since there are no other nodes joining the network, the mobile is able to hold onto these resources until other mobile nodes begin joining around 575s, as seen in Figure 6.20.

In reviewing the advertised network scores perceived by the mobile nodes joining the network, in Figures 6.22, 6.23, and 6.24, it is clear that the behavior is aligning to the number of handovers permitted, with the prediction model perceived score values dipping

below the aggregate counterparts at around 410s. This decreased view of the network performance contributes to fewer mobile nodes deciding to join the network.

In the final wave, occurring between 577 and 590 seconds, as a result of the prior period allowing fewer mobile nodes into the network, the initial handover request scores seen by the prediction model are higher. In particular, the peak at 584s is due to several nodes joining the network concurrently and receiving a high QoS. Post joining, the QoS experienced by these nodes is diminished due to the contention over resources. This behavior is also consistent with what is seen in the sample score trend in Figure 6.20.

Figure 6.22: Scenario 1 comparison of the reputation scores perceived by the mobile nodes upon entering the Wi-Fi network at 5km/h, between the prediction model (blue) and aggregate model (red); first period of arrival of mobile nodes
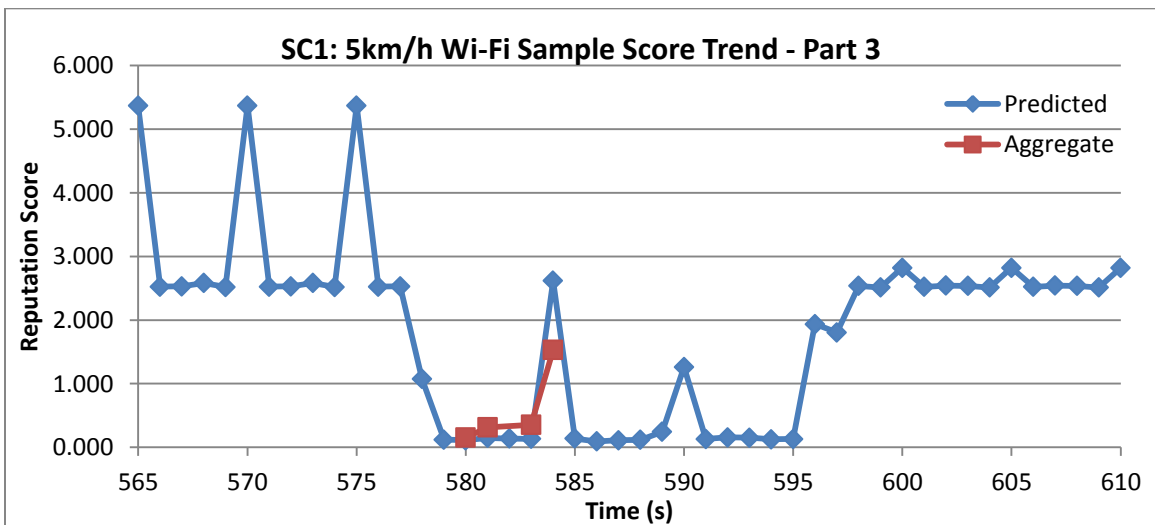
Figure 6.23: Scenario 1 comparison of the reputation scores perceived by the mobile nodes upon entering the Wi-Fi network at 5km/h, between the prediction model (blue) and aggregate model (red); Second period of arrival of mobile nodes

Figure 6.24: Scenario 1 comparison of the reputation scores perceived by the mobile nodes upon entering the Wi-Fi network at 5km/h, between the prediction model (blue) and aggregate model (red); third period of arrival of mobile nodes

6.2.3.2 Mobile Speed 40km/h:

Mobile nodes at this speed enter the Wi-Fi zone during two different periods, with the initial period being between 29s and 43s, and the second between 114s and 121s. During the initial phase, the congestion is not severe enough to impact the number of nodes allowed to join, so the same amount of handovers is seen in both the prediction and the aggregate method, as seen in Figure 6.25. During the second period an interesting observation at this speed is that in the Wi-Fi coverage area, more handovers occur in the

prediction model than in the aggregate model, as seen in Figure 6.26, contrary to what was expected.
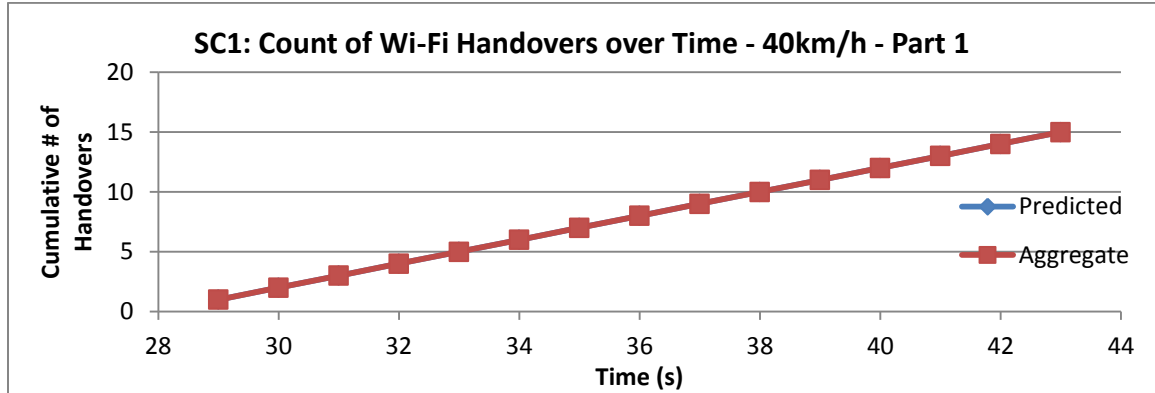
Figure 6.25: Scenario 1 comparison of the cumulative # of handovers over time between the proposed prediction model (blue) and the existing aggregate model (red) at 40km/h in the Wi-Fi network; first period of arrival of mobile nodes
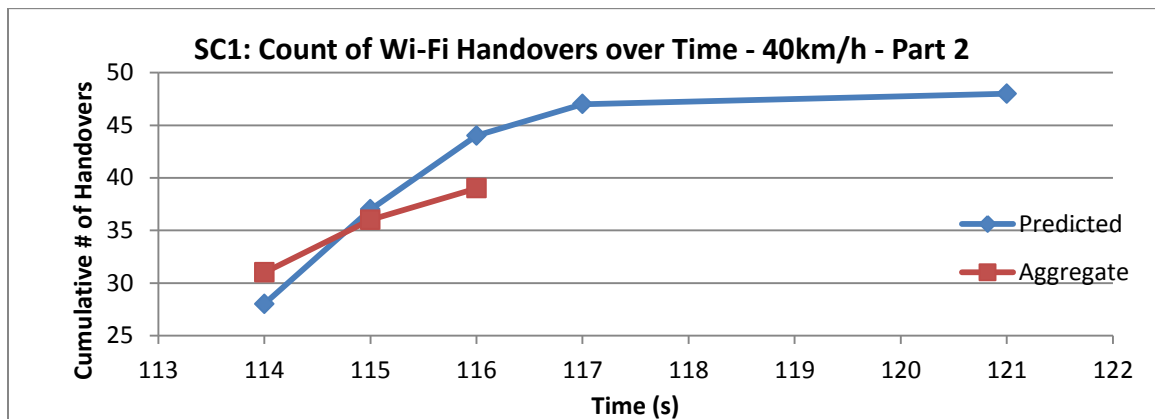
Figure 6.26: Scenario 1 comparison of the cumulative # of handovers over time between the proposed prediction model (blue) and the existing aggregate model (red) at 40km/h in the Wi-Fi network; second period of arrival of mobile nodes

This can be explained by the fact that when handovers were occurring into WiMAX, a significantly larger amount of nodes were allowed into WiMAX under the aggregate method, 44, whereas the prediction model only allowed 11. Furthermore, the last score calculated by the aggregate method during the WiMAX handovers was approximately 0.2, as seen in Figure 6.29. As a result, as the network becomes congested

around 115s, the nodes in the WiMAX network perceive their network as superior to the Wi-Fi and opt not to handover. In the prediction model, since numerous nodes did not handover to WiMAX initially, their observed QoS in UMTS is at 0.08, which permits more comparisons to result in a favorable view of the Wi-Fi network.
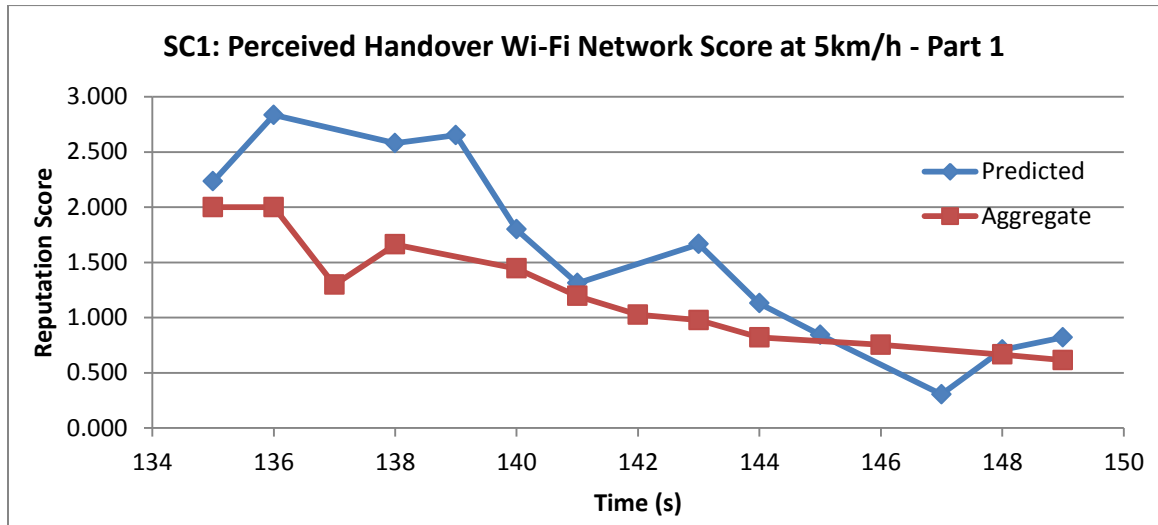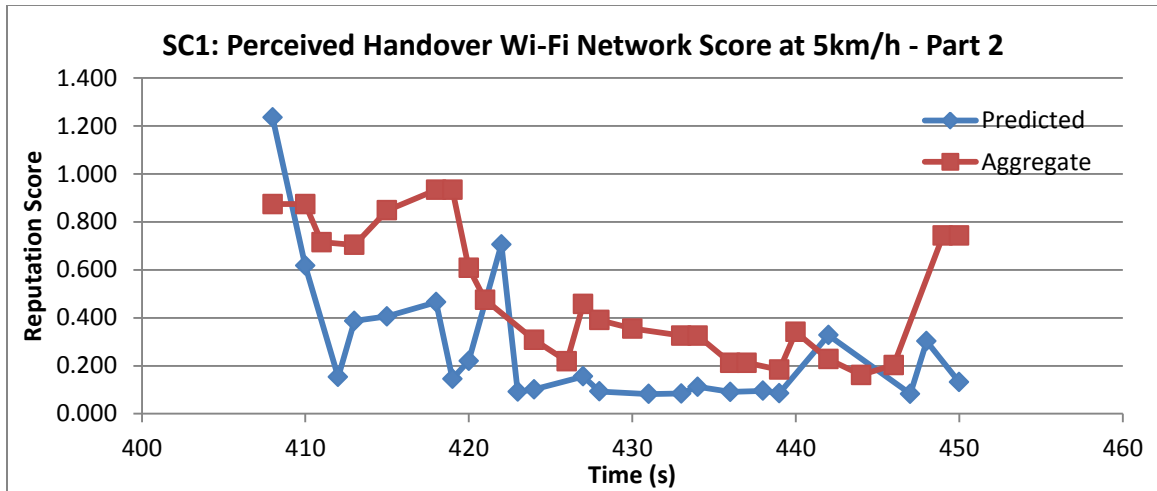

Figure 6.27: Scenario 1 comparison of the reputation scores sampled by the mobile nodes while within the Wi-Fi network at 40km/h, between the prediction model (blue) and aggregate model (red); first period of arrival of mobile nodes
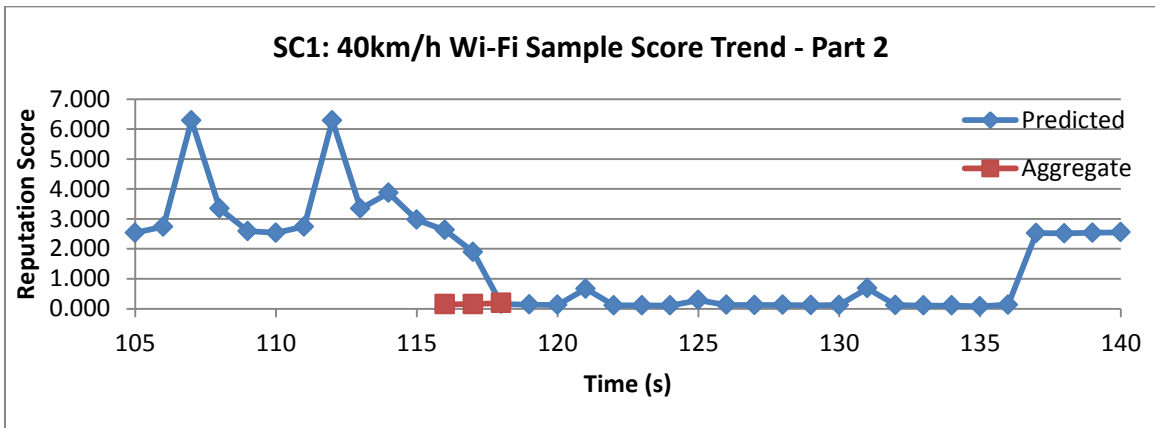
Figure 6.28: Scenario 1 comparison of the reputation scores sampled by the mobile nodes while within the Wi-Fi network at 40km/h, between the prediction model (blue) and aggregate model (red); second period of arrival of mobile nodes
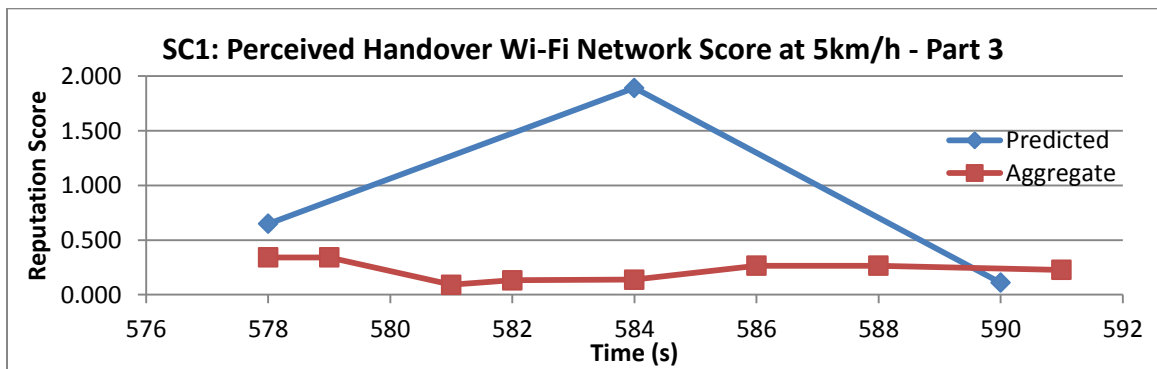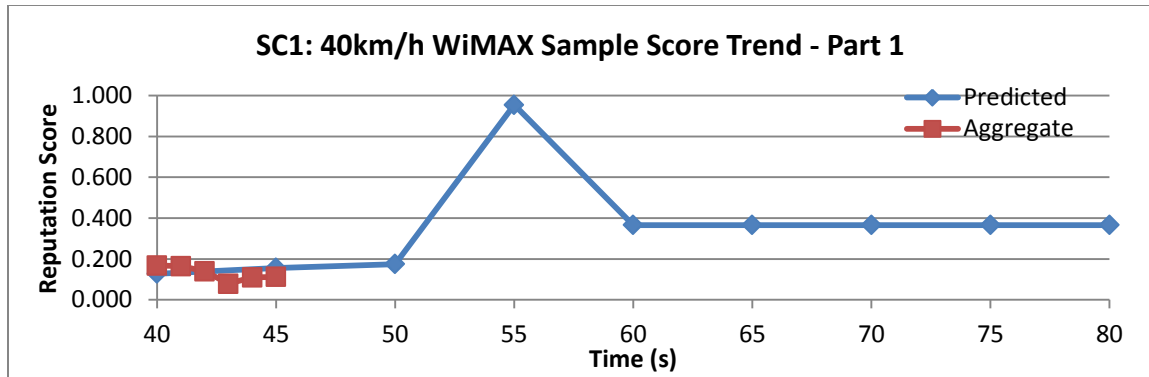
Figure 6.29: Scenario 1 comparison of the reputation scores sampled by the mobile nodes while within the WiMAX network at 40km/h, between the two models; during this period the first wave of mobile nodes are presented with the option to handover to Wi-Fi

Finally, in looking at the network scores perceived by the mobile nodes in the handover process in Figures 6.30 and 6.31, it is clear that the higher values seen in the prediction model align to the overall picture seen in the number of handovers and sample scores. The initial rise in the predicted score advertised by the prediction model is due to the fact that the predicted scores are significantly higher than 2 at the beginning of the period, as the nodes are experiencing a very good QoS. As congestion increases this score decreases. The aggregate calculation method on the other hand advertises the default network score initially due to mobile node sampling occurring 1 second after joining the network. Furthermore, due to the one second delay, the result is that there is slightly more congestion on the initial calculation, and as a result a lower initial score. Ultimately though it can be seen that both perceived handover scores follow each other, and this aligns to the fact that the number of handovers is the same over the first interval.
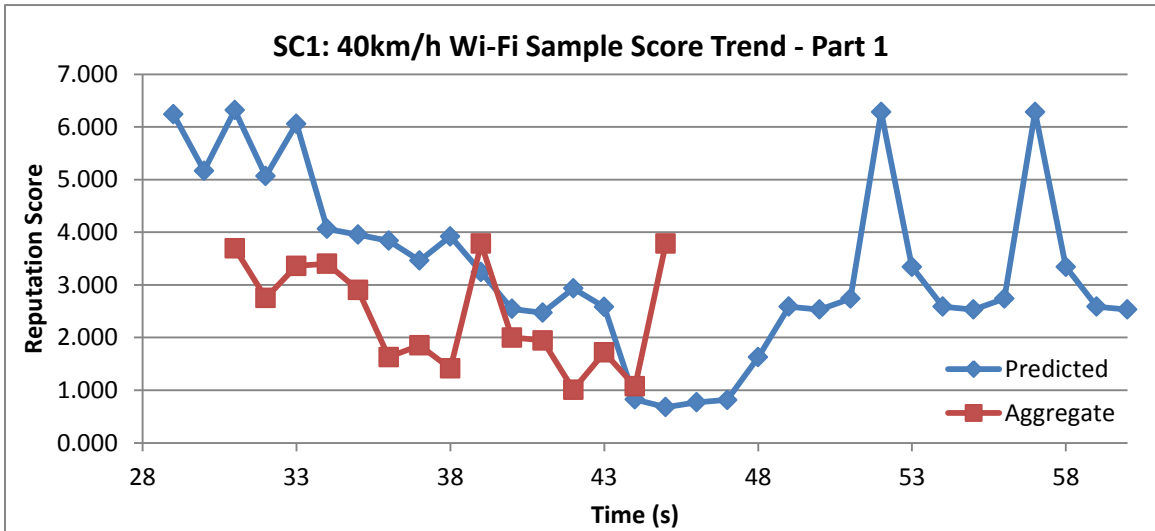
Figure 6.30: Scenario 1 comparison of the reputation scores perceived by the mobile nodes upon entering the Wi-Fi network at 40km/h, between the prediction model (blue) and aggregate model (red); first period of arrival of mobile nodes

Figure 6.31: Scenario 1 comparison of the reputation scores perceived by the mobile nodes upon entering the Wi-Fi network at 40km/h, between the prediction model (blue) and aggregate model (red); second period of arrival of mobile nodes

6.2.3.3 Mobile Speed 75km/h

Similar to the other two speeds, the first period where handovers take place in Wi-Fi yields the same number of handovers when running the simulation with the prediction model or the aggregate model, as seen in Figure 6.32. Subsequently, in the second period between 89s and 108s, the cumulative number of handovers is less in the prediction model than in the aggregate model, as expected, as seen in Figure 6.33.
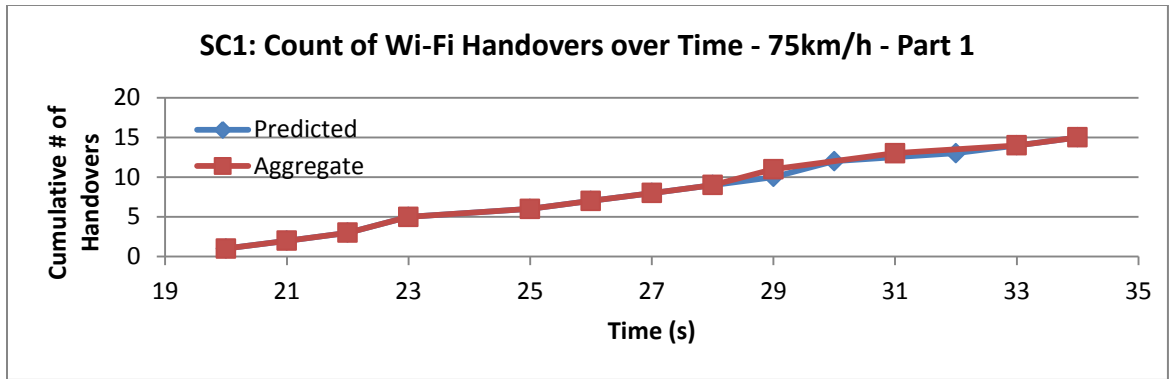
Figure 6.32: Scenario 1 comparison of the cumulative # of handovers over time between the proposed prediction model (blue) and the existing aggregate model (red) at 75km/h in the Wi-Fi network; first period of arrival of mobile nodes
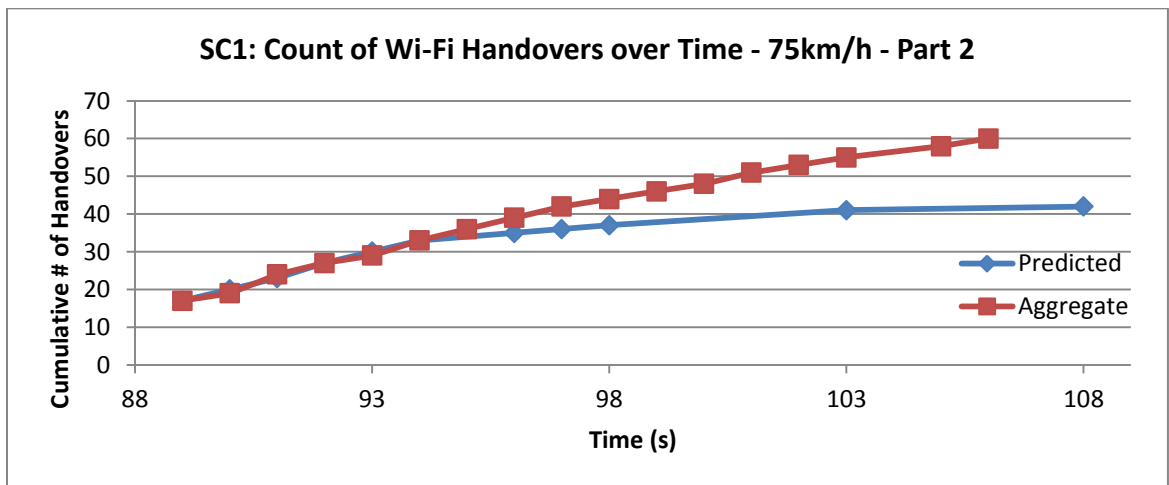
Figure 6.33: Scenario 1 comparison of the cumulative # of handovers over time between the proposed prediction model (blue) and the existing aggregate model (red) at 75km/h in the Wi-Fi network; second period of arrival of mobile nodes

The resulting fewer handovers in the prediction model is supported by the sample scores, which are lower in the prediction model after 99s, as seen in Figure 6.35. In Figure 6.34, the reason for the spike in sample scores around 29s in the prediction model is due to the fact that one of the nodes managed to obtain a good QoS upon entry into the network; similar behavior is seen with other nodes at the other speeds as well. After the initial entry, with the advent of other nodes contesting for resources, the QoS diminishes to the average experienced by the group.
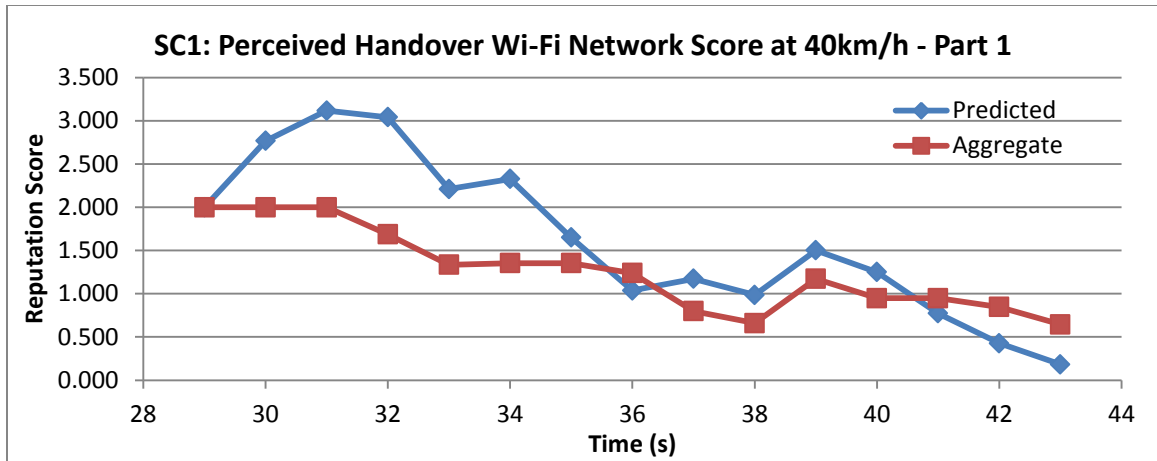
Figure 6.34: Scenario 1 comparison of the reputation scores sampled by the mobile nodes while within the Wi-Fi network at 75km/h, between the prediction model (blue) and aggregate model (red); first period of arrival of mobile nodes
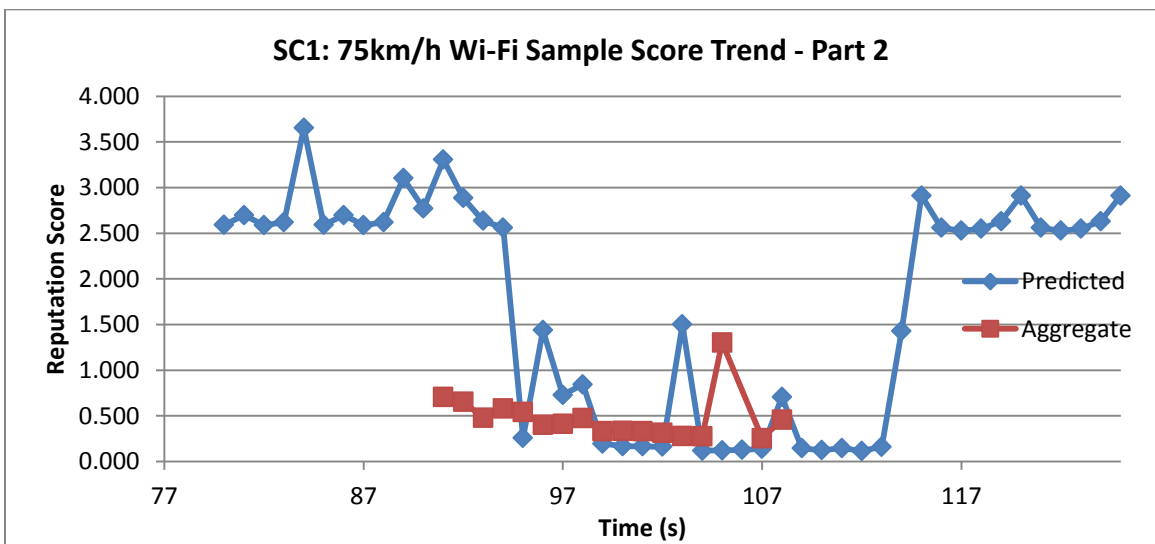
Figure 6.35: Scenario 1 comparison of the reputation scores sampled by the mobile nodes while within the Wi-Fi network at 75km/h, between the prediction model (blue) and aggregate model (red); second period of arrival of mobile nodes
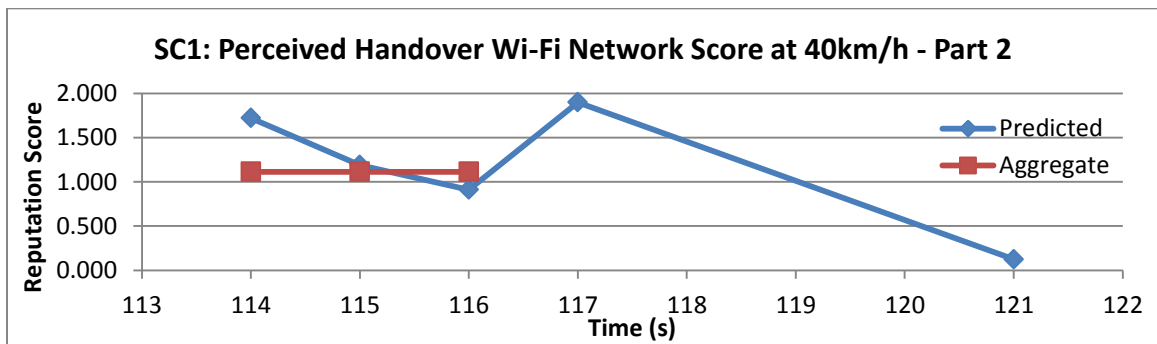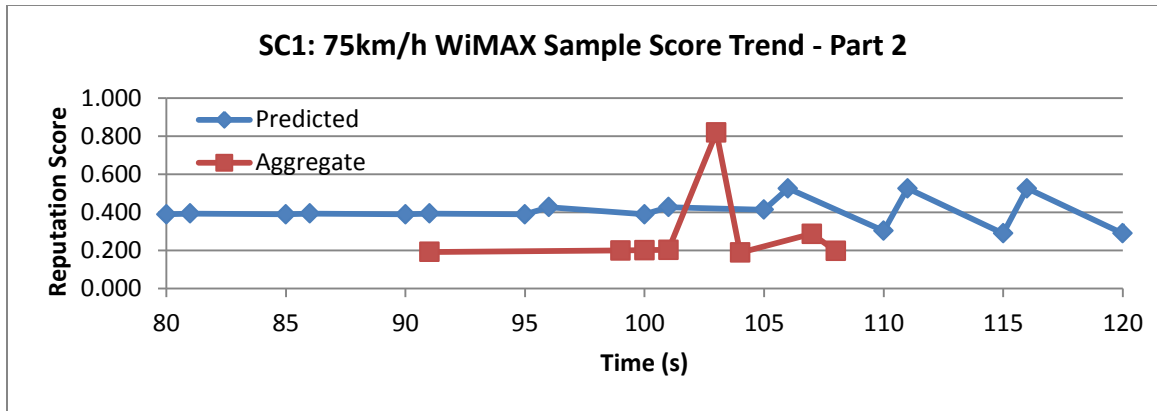
Figure 6.36: Scenario 1 comparison of the reputation scores sampled by the mobile nodes while within the WiMAX network at 75km/h, between the two models; during this period the second wave of mobile nodes are presented with the option to handover to Wi-Fi

In the second period of handovers between 89 and 108 seconds, during the peak congestion time, it can be also seen from Figure 6.36 that even though fewer handovers are facilitated by the prediction model into Wi-Fi, the mobile nodes remaining in the WiMAX network experience better service than their aggregate counterparts. Furthermore, the mobile nodes in the prediction model in the Wi-Fi coverage also stabilize at a very favorable value as well, as seen in Figure 6.35, indicating overall that the model has performed well in balancing both networks. Note that the one spike in Figure 6.36 in the aggregate model at 103s is due to one mobile node obtaining a good QoS upon entry into the network.

In reviewing the network advertised scores perceived by the mobile nodes joining the network, it is clear that both are in a congestion trend during both periods of nodes joining the network. In both Figures 6.37 and 6.38, the initial rise in the prediction model is due to the mobile nodes experiencing a good QoS initially upon entry into the network. This service tapers off though, as is seen in the sample score figures, resulting in diminishing advertised scores.
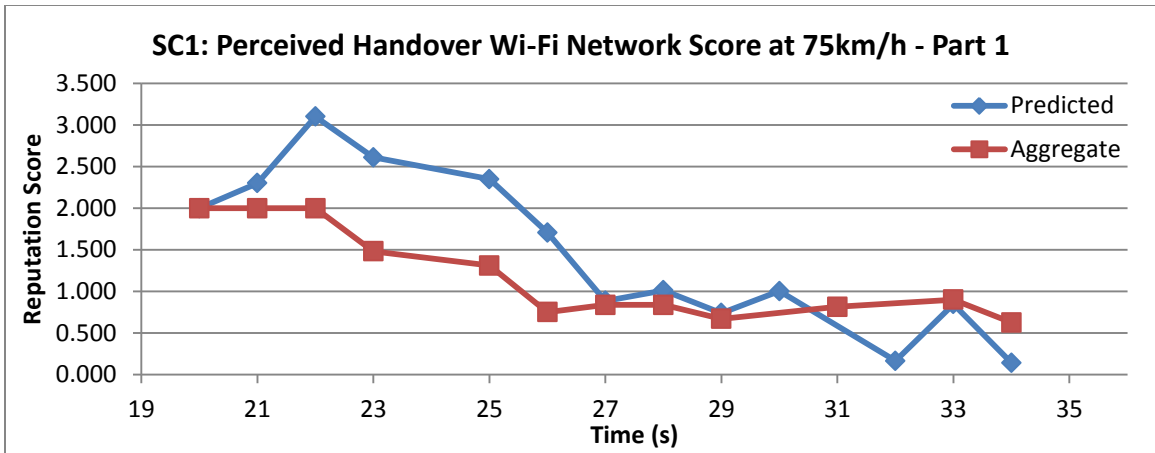
Figure 6.37: Scenario 1 comparison of the reputation scores perceived by the mobile nodes upon entering the Wi-Fi network at 75km/h, between the prediction model (blue) and aggregate model (red); first period of arrival of mobile nodes
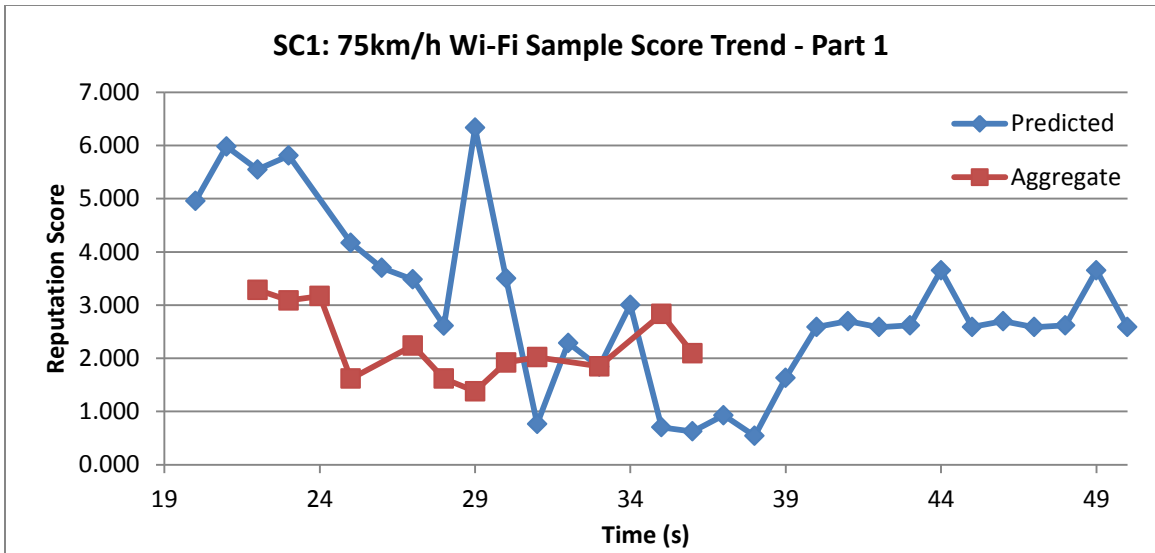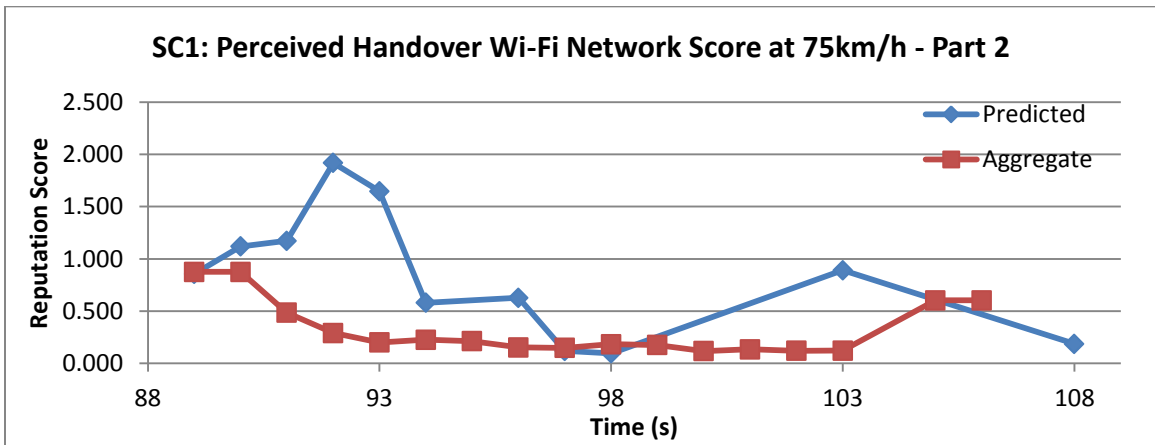

Figure 6.38: Scenario 1 comparison of the reputation scores perceived by the mobile nodes upon entering the Wi-Fi network at 75km/h, between the prediction model (blue) and aggregate model (red); second period of arrival of mobile nodes

6.3   Multi-Node Observations: Increasing & Decreasing Load

The second example simulated in order to assess the performance of the proposed model, involves a more slowly increasing load on a network, followed by a decrease in load on a network. This attempts to assess how the model will behave when multiple nodes are leaving a network while others are attempting to join.

6.3.1 Scenario Description

In this scenario, mobile nodes initially move into the WiMAX coverage area at a slower rate than in example 1 (less overlap between groups), then progress to the Wi-Fi, and finally several of the groups after having reached the Wi-Fi coverage area begin to exit back into WiMAX. Similarly to the initial scenario, one could assume that due to the gradually increasing load in WiMAX, the prediction method would reduce the number of handovers facilitated, and during the slowdown of traffic in Wi-Fi, there would be more handovers permitted. But, as seen in the results, it is in fact the opposite which happens. The scenario in question plays out as follows:

1. At 10s the group below the WiMAX BS begins moving into the WiMAX coverage area, with one node beginning to move every second.

2. Similarly, at 26s, 42s, and 60s, the groups above, to the right, and to the left of the WiMAX BS begin moving into WiMAX network respectively.

3. Subsequently, at 60s, a mobile node from the group below the WiMAX BS begins moving towards the Wi-Fi BS, with one node beginning the trek every 2.5 seconds.

4. At 80s and 93s the mobile nodes from the groups above and to the right of the WiMAX BS respectively begin moving towards Wi-Fi.

5. Finally, at 145s and 160s the mobile nodes from the groups to the left and above the WiMAX BS respectively begin *leaving* the Wi-Fi group, causing a decreasing load in the Wi-Fi network.

6. The simulation ends after 250s.

This scenario is repeated for the mobile speeds of 75km/h and 5km/h, and in order to ensure that the scenario is the same, the timings are adjusted to compensate for the speed change. For instance, in the case of 5km/h, the overall simulation time is 850s instead of 250s.
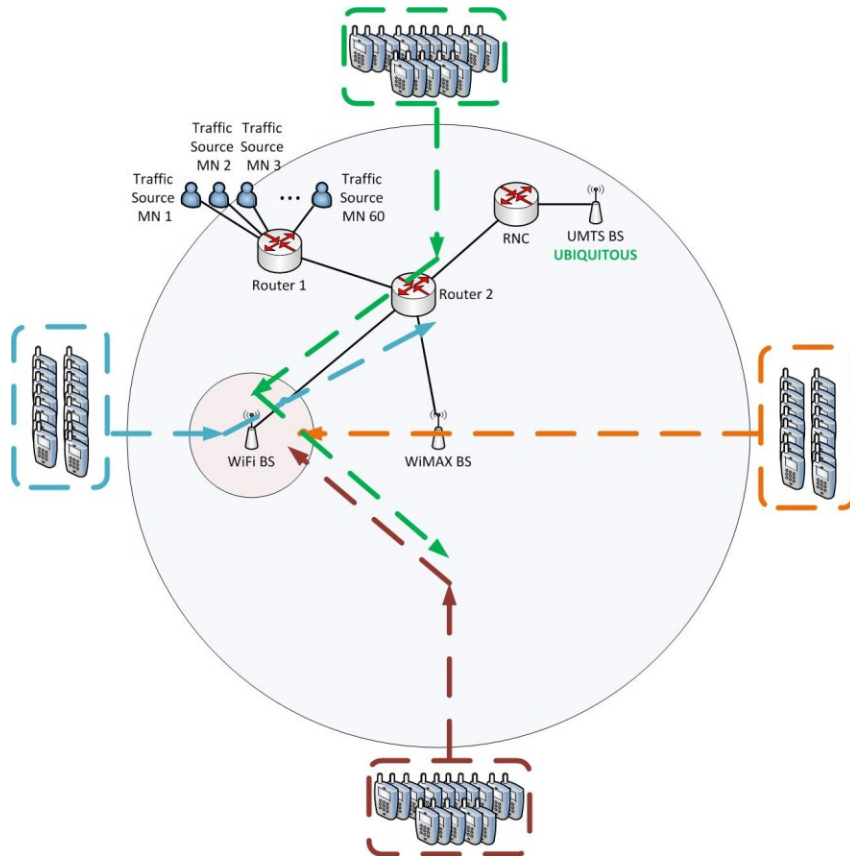
Figure 6.39: Topology and mobile node path for WiMAX and WiFi with increasing and decreasing traffic scenario

6.3.2 WiMAX Results

In this section, the results of the three different speeds are reviewed as they pertain to the WiMAX coverage area. In general, the comparison between the total number of handovers realized upon completion of the simulation within the proposed

model and original model are different than what was initially thought, as can be seen in Figure 6.40.

SC2: Number of Handovers in WiMAX
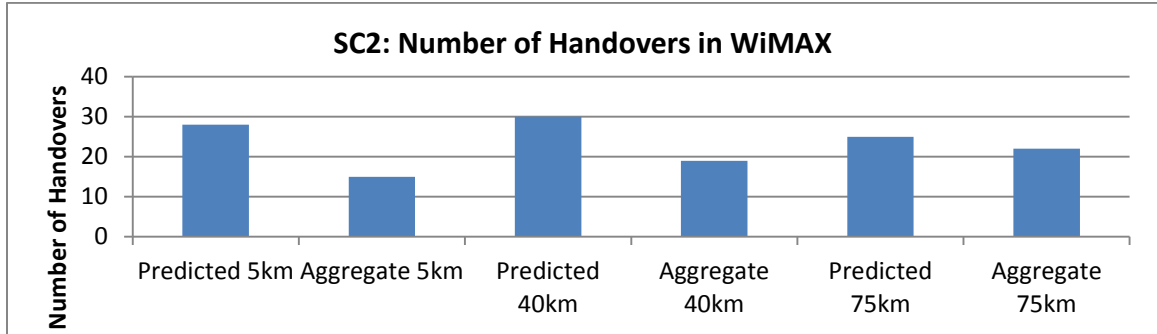
Figure 6.40: Scenario 2 comparison of the total number of VHOs between the proposed prediction model and the existing aggregate model at mobile node speeds of 5km/h, 40km/h, and 75km/h in WiMAX

6.3.2.1 Mobile Speed 5km/h:

In this scenario, due to the location of the groups and when they begin moving toward the WiMAX network, multiple nodes reach the coverage area at the same time every 5 seconds, beginning at 55s. The difference in this scenario from scenario 1 is that number of nodes arriving at the coverage area is less in scenario 2, since in this scenario each group successfully reaches the WiMAX coverage area prior to the next group arriving. In scenario 1, there was overlap between the groups arriving at the WiMAX coverage area. As can be seen in Figure 6.41, the number of handovers permitted over time is greater in the prediction model than in the aggregate.
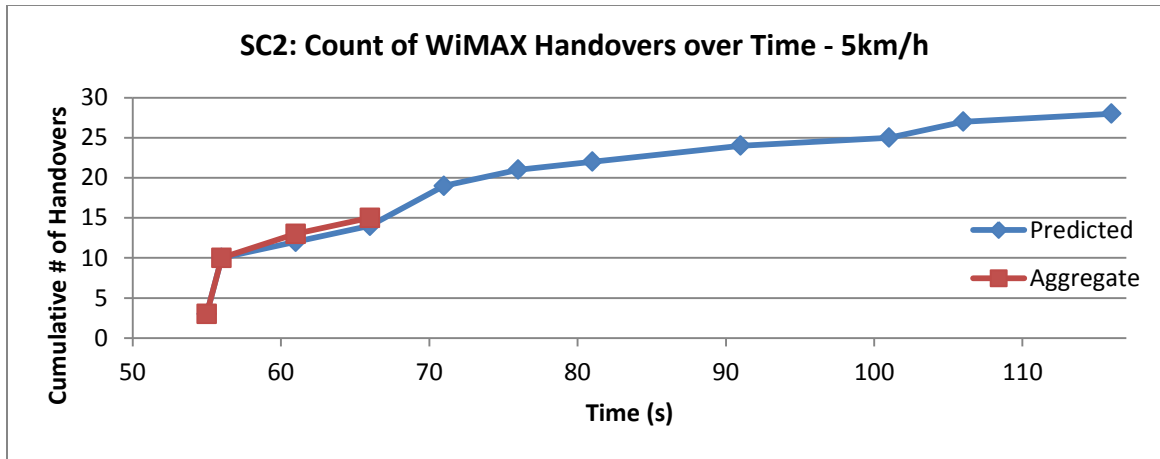
Figure 6.41: Scenario 2 comparison of the cumulative # of handovers over time between the proposed prediction model (blue) and the existing aggregate model (red) at 5km/h in the WiMAX network

When a group of nodes arrive at the coverage area, they arrive approximately concurrently, i.e. if 10 nodes arrive at 55s, several nodes arrive within 0.3s of each other. In the prediction model, the nodes calculate the sample scores immediately upon entry and as a result, the score calculated gradually takes into account the full group which arrived approximately concurrently. In contrast, the aggregate method samples 1s after joining the network, resulting in a poorer score calculated by each node since by this time the entire group of nodes which arrived at the coverage area at the same time have all begun operating in the network. Furthermore, since the mobile nodes are joining the network much more gradually than in scenario 1, the degradation of the network is not as steep, permitting the predicted score to stay high enough to permit more handovers. Consequently, there are more handovers facilitated via the prediction model than in the aggregate model.

This behavior is reflected in the fact that the sample scores calculated by the mobile nodes in the prediction model are higher than the aggregate, as seen in Figure 6.42. Also, in reviewing the advertised network scores perceived by the mobile nodes

entering the network, depicted in Figure 6.43, it is evident that at 66s, when the aggregate model hits its lowest advertised score, the prediction model begins to increase slightly, providing a better perceived score which permits more handovers.

Figure 6.42: Scenario 2 comparison of the reputation scores sampled by the mobile nodes while within the WiMAX network at 5km/h, between the prediction model (blue) and aggregate model (red)
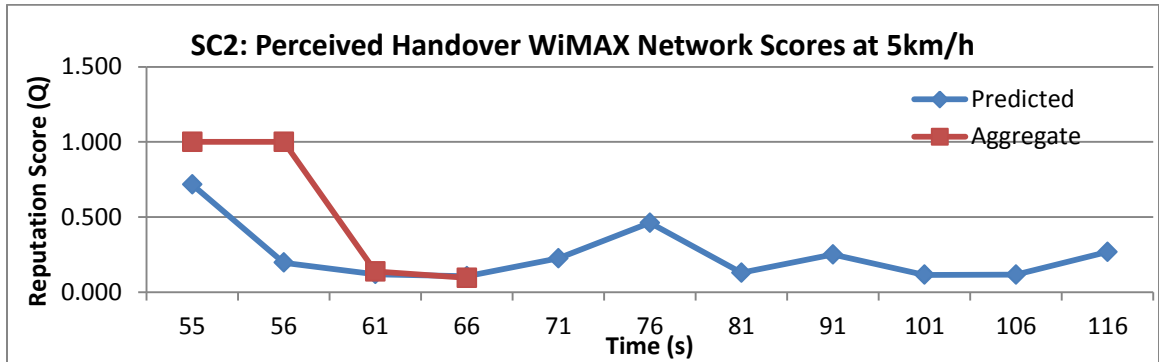
Figure 6.43: Scenario 2 comparison of the reputation scores perceived by the mobile nodes upon entering the WiMAX network at 5km/h, between the prediction model (blue) and aggregate model (red)

6.3.2.2 Mobile Speed 40km/h:

By increasing the speed to 40km/h, the same behavior is observed as in the case where the speed was 5km/h, with virtually the same number of handovers occurring in the prediction model under both speeds (28 at 5km/h, 30 at 40km/h). In the aggregate

96

model, 4 more handovers are experienced at the higher speed (19 vs. 15). This is observed to be due to the fact that fewer nodes initially arrive concurrently at the WiMAX network, resulting in a better initial score calculated 1s after entry. Since the initially score calculated after 1s is higher than in the 5km/h scenario, it takes more poor calculated scores to degrade the aggregate score below the threshold. The cumulative amount of handovers over time can be seen in Figure 6.44.

Figure 6.44: Scenario 2 comparison of the cumulative # of handovers over time between the proposed prediction model (blue) and the existing aggregate model (red) at 40km/h in the WiMAX network

In reviewing the sample scores taken by the mobile nodes in the network, it is evident that the behavior supports the above discussion in the sense that the scores calculated by the prediction model are higher than the aggregate, as seen in Figure 6.46. Furthermore, these higher scores in the prediction model contribute to a network advertised score that is also higher in the prediction model, as seen in Figure 6.45 after 30s. This corresponds to where the sample scores start becoming higher in the prediction model as well.

Figure 6.45: Scenario 2 comparison of the reputation scores perceived by the mobile nodes upon entering the WiMAX network at 40km/h, between the prediction model (blue) and aggregate model (red)
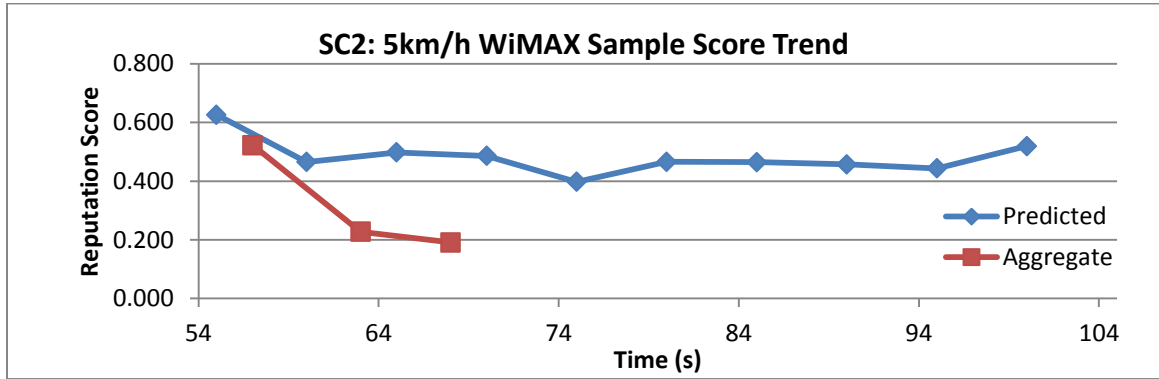

Figure 6.46: Scenario 2 comparison of the reputation scores sampled by the mobile nodes while within the WiMAX network at 40km/h, between the prediction model (blue) and aggregate model (red)

6.3.2.3 Mobile Speed 75km/h:

At 75km/h, similar to the other speeds, there are still more handovers observed in the prediction model than in the aggregate mode (25 to 22), as can be seen in Figure 6.47.
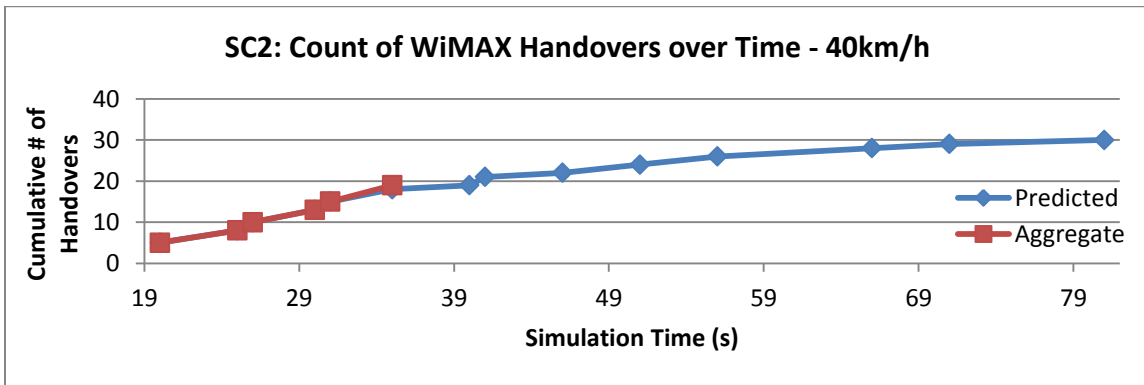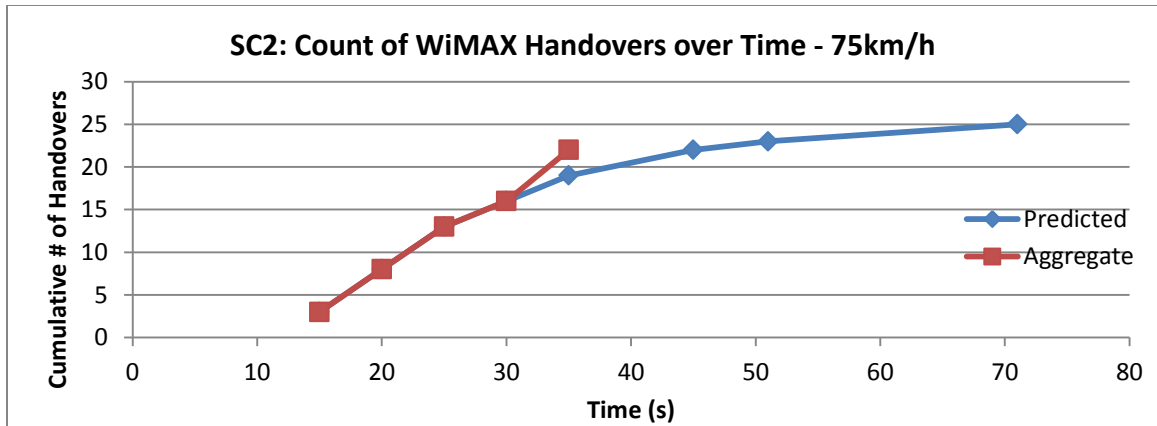
Figure 6.47: Scenario 2 comparison of the cumulative # of handovers over time between the proposed prediction model (blue) and the existing aggregate model (red) at 75km/h in the WiMAX network

In observing the advertised network scores perceived by the mobile nodes, depicted in Figure 6.48, it can be noted that there is a peak experienced at 30s that tapers off as more nodes join the network. This is due to the fact that 4 nodes from the red group are significantly closer to the BS than the other nodes at this speed, and as such experience better QoS. This is reflected in the scores they sample, seen in Figure 6.49, which boosts the overall aggregate score calculated by the network agent. Although they continue to experience better scores than the rest, this diminishes as more nodes join the network, as can be seen in the time frame after 30s.
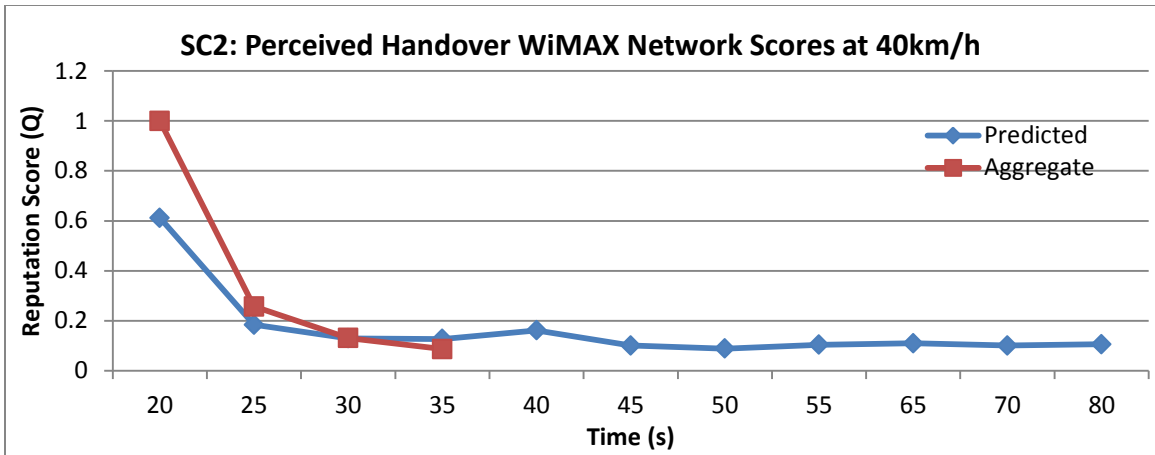

Figure 6.48: Scenario 2 comparison of the reputation scores perceived by the mobile nodes upon entering the WiMAX network at 75km/h, between the prediction model (blue) and aggregate model (red)
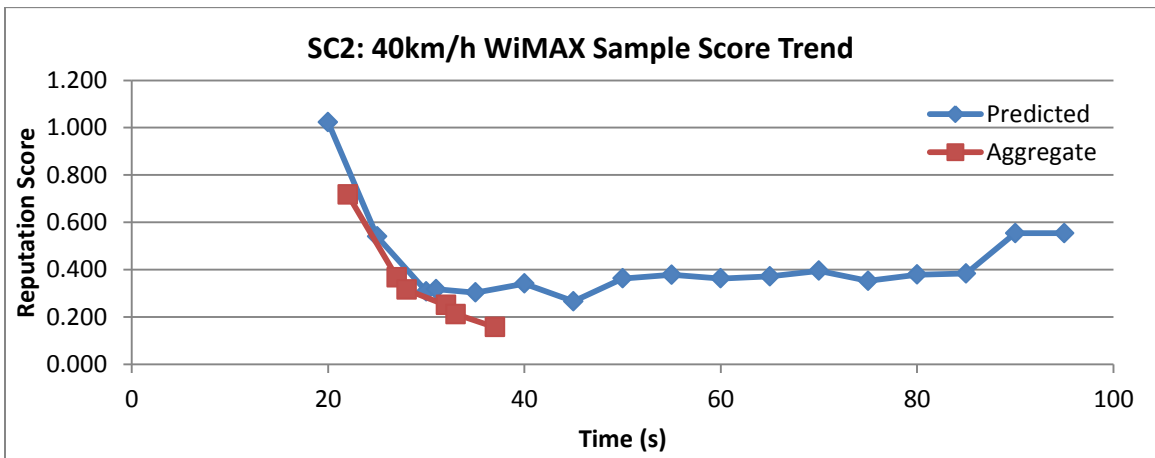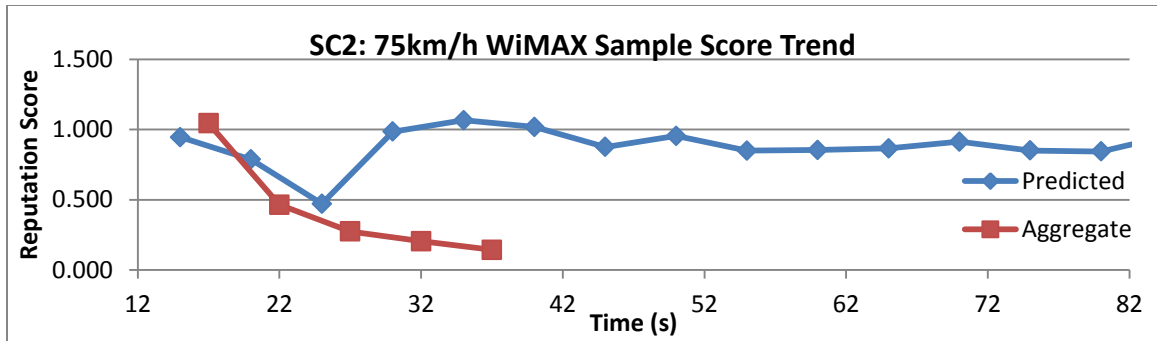
Figure 6.49: Scenario 2 comparison of the reputation scores sampled by the mobile nodes while within the WiMAX network at 75km/h, between the prediction model (blue) and aggregate model (red)

6.3.3 Wi-Fi Results

Based on the first scenario, the initial assumption is that the mobile nodes entering a Wi-Fi network where nodes are also leaving the coverage will result in more handovers facilitated through the prediction model. In fact, the reverse is true as seen in Figure 6.50, due to the fact that the prediction model requires a longer series of good results to boost the predicted score. Furthermore, as will be seen in this section, although some mobile nodes do begin to experience good service when traffic begins to diminish, these results are dampened by the other continuously polling nodes that are experiencing poorer QoS.
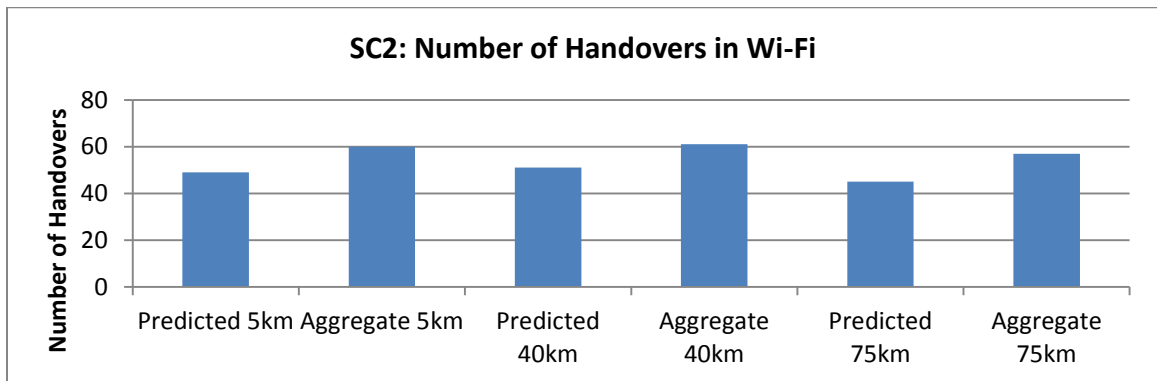
Figure 6.50: Scenario 2 comparison of the total number of VHOs between the proposed prediction model and the existing aggregate model at mobile node speeds of 5, 40, and 75 km/h in Wi-Fi

6.3.3.1 Mobile Speed 5km/h

Due to the slow speed and distance between the groups, the mobile nodes arrive at the Wi-Fi coverage area during three different periods. During the first period of arrivals, all the mobile nodes experience good quality of service and as a result there is no difference in the number of handovers between the two models, as can be seen in Figure 6.51.
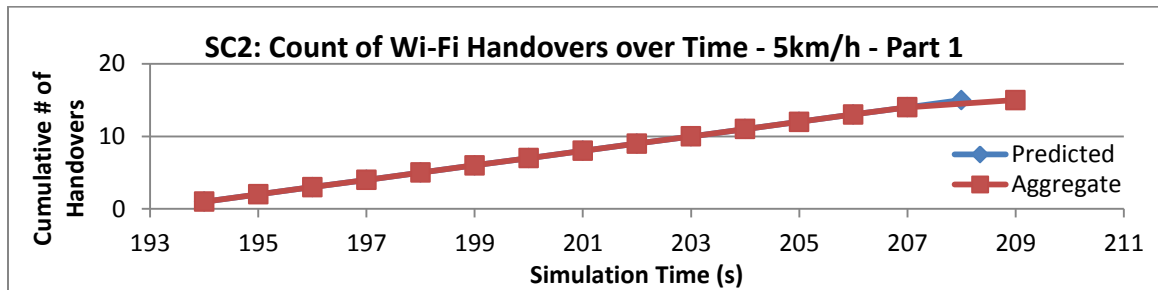
Figure 6.51: Scenario 2 comparison of the cumulative # of handovers over time between the proposed prediction model (blue) and the existing aggregate model (red) at 5km/h in the Wi-Fi network; first period of arrival of mobile nodes

In observing the associated sample scores calculated during this period in Figure 6.52, and the network advertised scores in Figure 6.53, it can be seen that both the aggregate and prediction methods are heading in a decreasing slope over time, but generally good scores are still maintained in both models within the first period.

Figure 6.52: Scenario 2 comparison of the reputation scores sampled by the mobile nodes while within the Wi-Fi network at 5km/h, between the prediction model (blue) and aggregate model (red); first period of arrival of mobile nodes
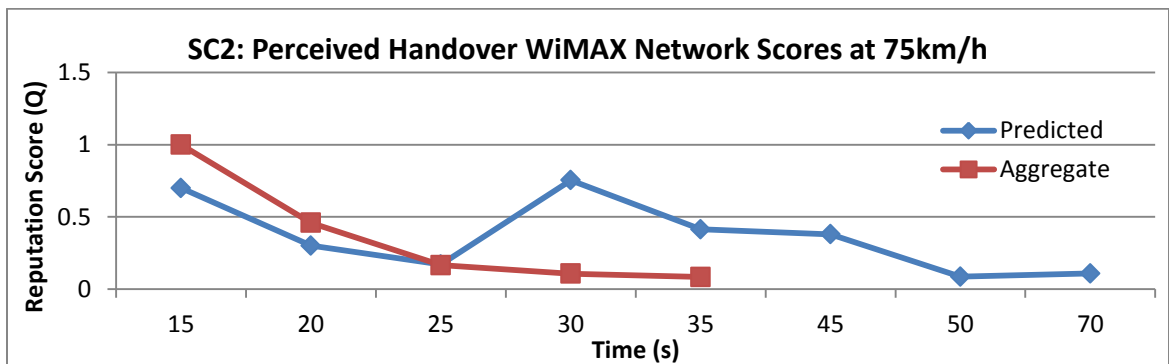
Figure 6.53: Scenario 2 comparison of the reputation scores perceived by the mobile nodes upon entering the Wi-Fi network at 5km/h, between the prediction model (blue) and aggregate model (red); first period of arrival of mobile nodes

During the second period of mobile nodes joining the Wi-Fi network, between 500 and 700 seconds, there are three more handovers experienced in the aggregate model, 30, compared to the prediction model, 27, as seen in Figure 6.54.

Figure 6.54: Scenario 2 comparison of the cumulative # of handovers over time between the proposed prediction model (blue) and the existing aggregate model (red) at 5km/h in the Wi-Fi network; second period of arrival of mobile nodes

In observing the sample scores in the WiMAX and Wi-Fi areas in Figures 6.55 and 6.56 respectively, it can be seen that due to the numerous handovers originally in WiMAX under the prediction model, the prediction model sample scores are averaging

around 0.5 in WiMAX, whereas the aggregate model which had significantly less handovers and only calculated the score upon entry, is around 1.

Figure 6.55: Scenario 2 comparison of the reputation scores sampled by the mobile nodes while within the WiMAX network at 5km/h, between the two models; during this period the second wave of mobile nodes are presented with the option to handover to Wi-Fi
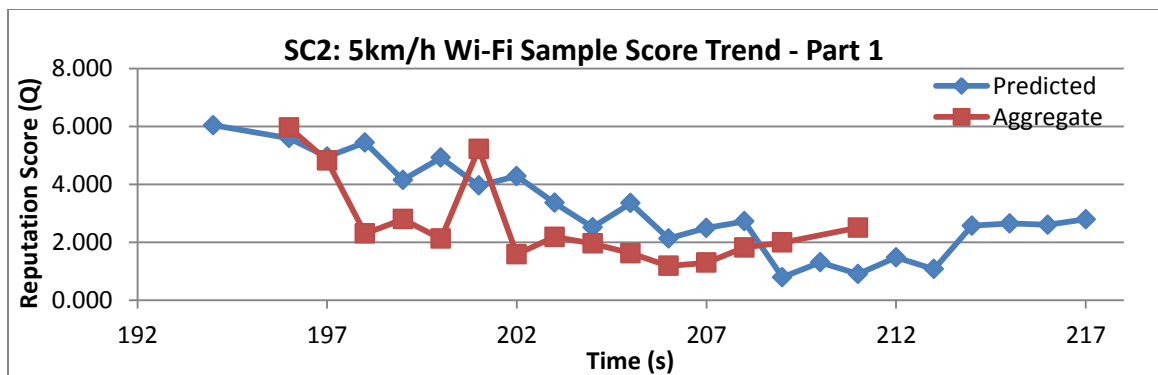

Figure 6.56: Scenario 2 comparison of the reputation scores sampled by the mobile nodes while within the Wi-Fi network at 5km/h, between the prediction model (blue) and aggregate model (red); second period of arrival of mobile nodes

As a result, when the mobile nodes begin joining the Wi-Fi network, higher congestion begins occurring and some of the predicted scores advertised by the network agent are below 0.5, as seen in Figure 6.57, resulting in some of the mobile nodes in the prediction model choosing to stay in the WiMAX coverage area.

The oscillations seen in Figures 6.56 and 6.57 are due to the fact that all the nodes are typically arriving at the Wi-Fi coverage area at unique times as a result of travelling a

long distance, as compared to travelling into WiMAX. This results in much less averaging of scores at each second when reporting in the above figures, and since the network access is contention based, there is some variance on the QoS obtained by each node. Typically the initial values of the scores are higher and begin to taper off, which also leads to the oscillatory behavior.
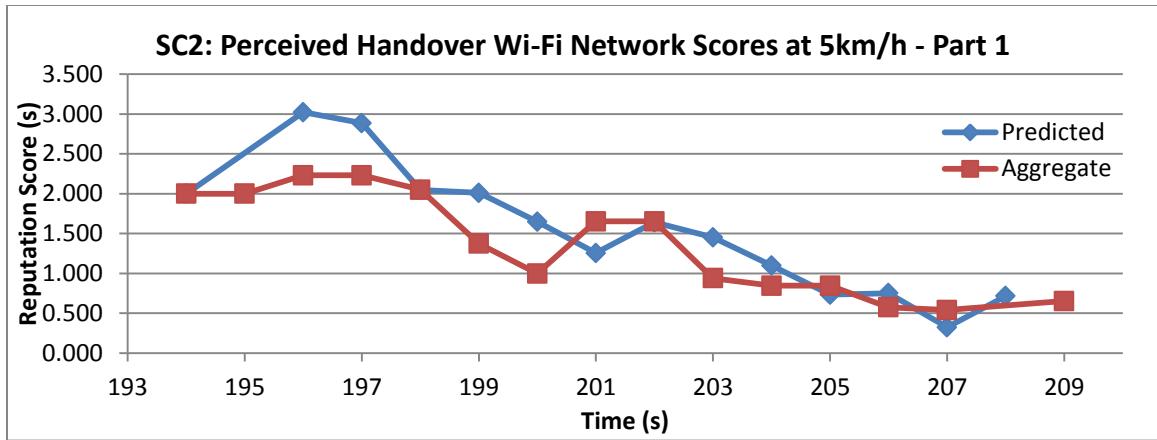

Figure 6.57: Scenario 2 comparison of the reputation scores perceived by the mobile nodes upon entering the Wi-Fi network at 5km/h, between the prediction model (blue) and aggregate model (red); second period of arrival of mobile nodes

In the final period of Wi-Fi arrivals around 730s, it can be seen in Figure 6.58 that significantly more handovers now occur in the aggregate model as compared to the prediction model.
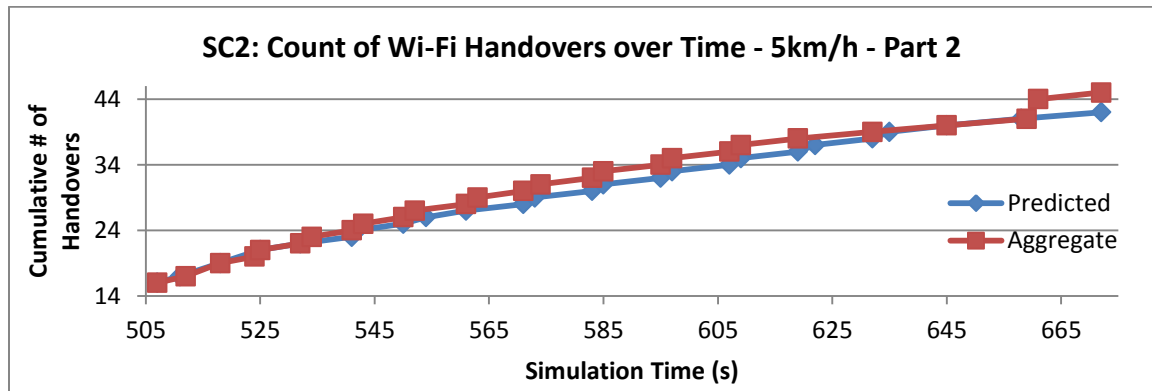
Figure 6.58: Scenario 2 comparison of the cumulative # of handovers over time between the proposed prediction model (blue) and the existing aggregate model (red) at 5km/h in the Wi-Fi network; third period of arrival of mobile nodes

During the period in which the third group arrives, the mobile nodes from two of the groups already in the Wi-Fi zone leave, sequentially. This combination of arrivals and departures between 730s and 760s also contributes to the oscillatory behavior during this timeframe in the sample scores, as seen in Figure 6.60, and the generally increasing scores advertised by the network agent for handovers (since there are more nodes leaving than entering) visible in Figure 6.59. As can be seen from the prediction method scores, the initial congestion drives the score down sharply, and since there is only a gradual ease of traffic, the sample scores slowly become better until 760s when the network stabilizes.
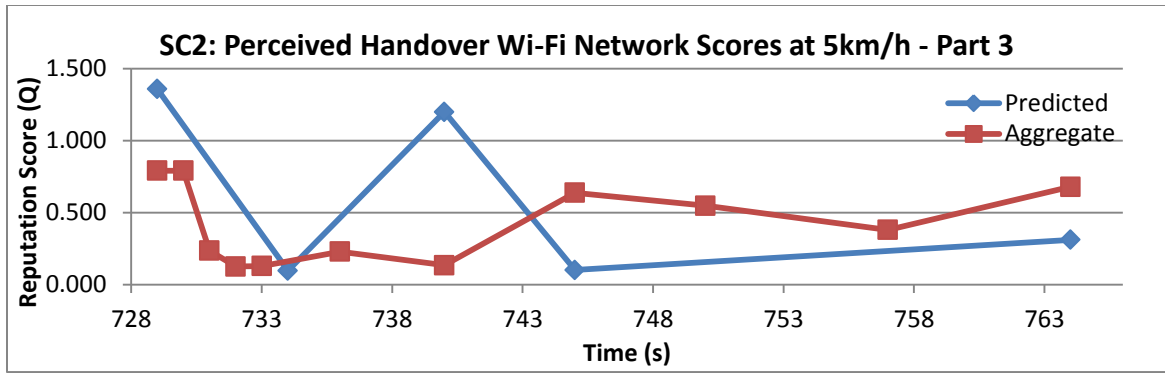
Figure 6.59: Scenario 2 comparison of the reputation scores perceived by the mobile nodes upon entering the Wi-Fi network at 5km/h, between the prediction model (blue) and aggregate model (red); third period of arrival of mobile nodes

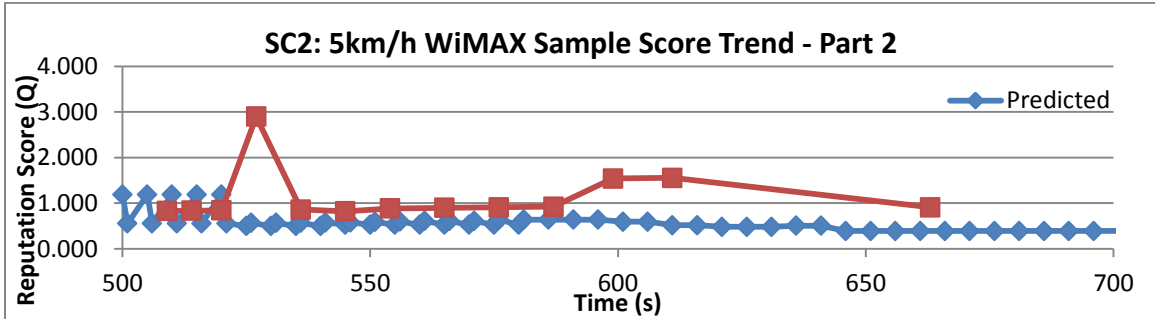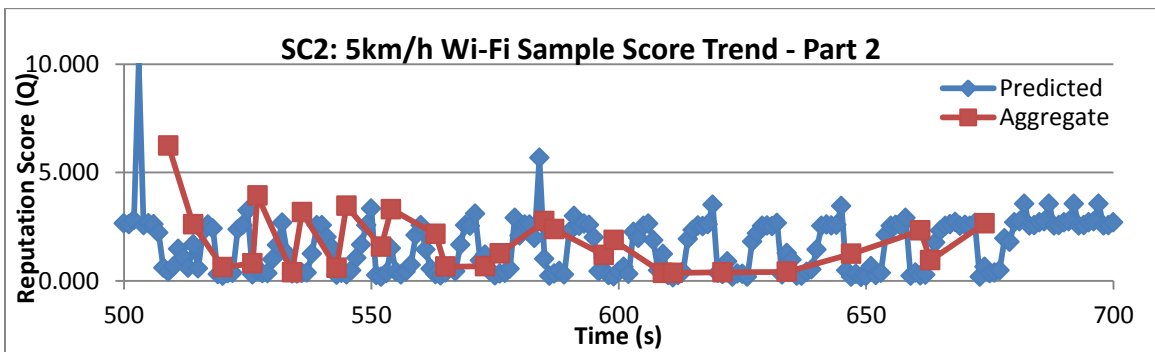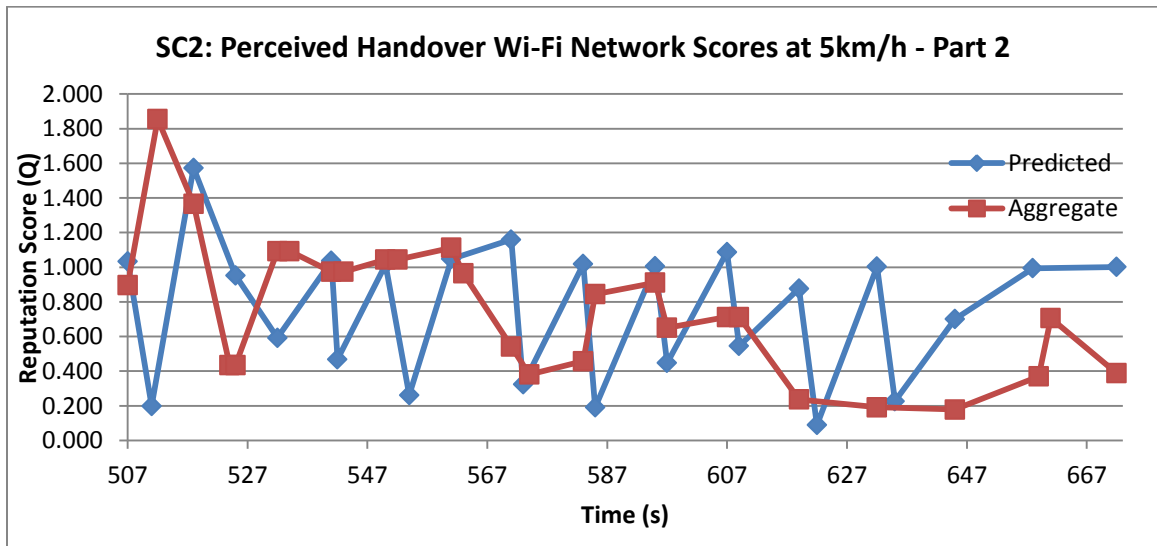**Figure 6.60: Scenario 2 comparison of the reputation scores sampled by the mobile nodes while within the Wi-Fi network at 5km/h, between the prediction model (blue) and aggregate model (red); third period of arrival of mobile nodes**

Furthermore, in looking at the scores in the prediction model, it can be noted that the spikes seen in the scores are nodes joining the network and initially experiencing good QoS, but which rapidly decreases to the average QoS experienced by the rest of the nodes. For example, the node joining at 740s experiences initially very low jitter and delay, and as a result, generates a high score. On the next sample at 745s, the same node scores the network lower, albeit still higher than the average. On the third sample, the same node is now experiencing the same congestion as the rest of the network. In contrast, in the aggregate model, the initial sample after joining the network is all that is

recorded by the node, so as a result, the list of sample scores seems higher than the prediction model, and subsequently the network advertises higher scores in the aggregate model.

Finally, in addition to the fact that the aggregate model produces higher network advertised scores due to the lack of sample data, fewer nodes joined the WiMAX coverage area in the aggregate model and therefore more have remained in UMTS, which has a flat QoS of 0.08. This means that it requires very little for the node to switch into the Wi-Fi zone. The prediction model on the other hand, had many nodes move into WiMAX, and up until 740s the WiMAX scores look better than the Wi-Fi, as seen in Figure 6.61, resulting in less handovers in the prediction model.

Figure 6.61: Scenario 2 comparison of the reputation scores sampled by the mobile nodes while within the WiMAX network at 5km/h, between the two models; during this period the third wave of mobile nodes are presented with the option to handover to Wi-Fi

6.3.3.2 Mobile Speed 40km/h:

When the speed is increased to 40km/h, and with the departure times of the nodes adjusted to match the speed, the nodes arrive at the Wi-Fi coverage area in 4 different

periods. In the first period, between 79 and 95 seconds, both the prediction model and the aggregate model experience the same number of handovers (16), due to the lack of congestion in the network, as seen in Figure 6.62.
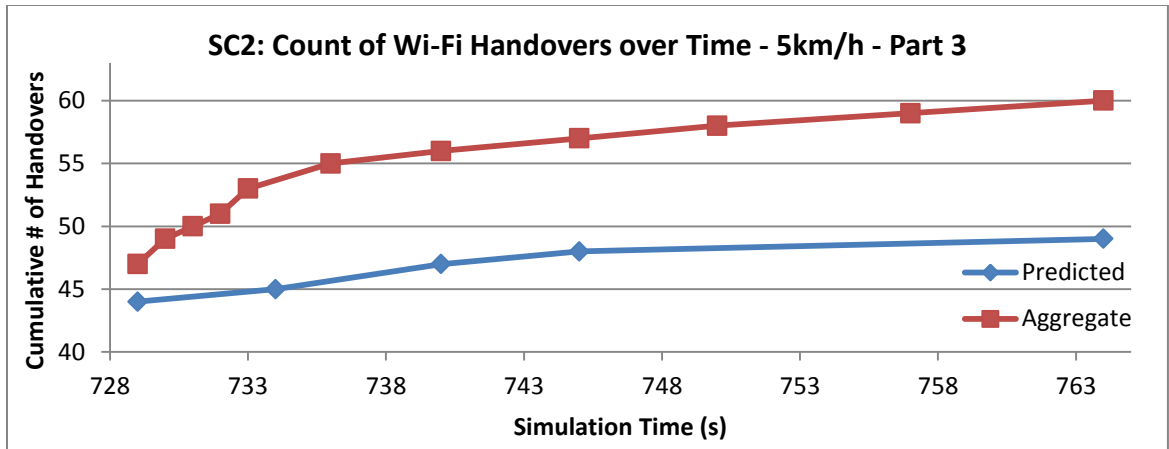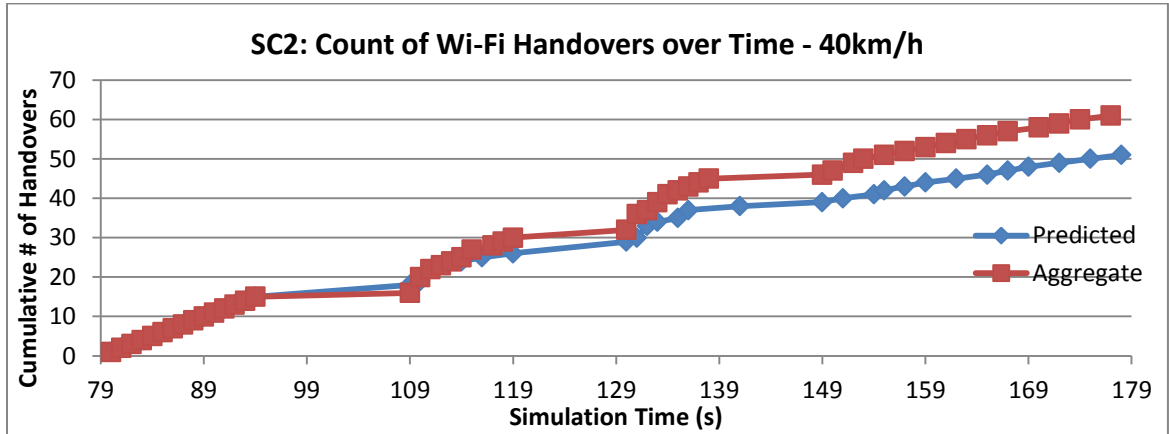

Figure 6.62: Scenario 2 comparison of the cumulative # of handovers over time between the proposed prediction model (blue) and the existing aggregate model (red) at 40km/h in the Wi-Fi network

This is expected since at this point the WiMAX network is still in congestion, with the sample scores residing around 0.5 (as seen in figure 6.65). In the Wi-Fi sample scores seen below (Figure 6.64), the initial rise in the predicted score is due to the fact that the nodes calculate the score upon entry. In the aggregate model, since the score is sampled one second after entry, there is more congestion in the network at this point. The peak in the prediction model at 90s is due to three successively increasing scores sampled between 87 and 89 seconds, and the peak in the aggregate model at 87 seconds is due to a node experiencing very low delay and jitter. These peaks also correspond to the rises seen in the network advertised scores in Figure 6.63.

Figure 6.63: Scenario 2 comparison of the reputation scores perceived by the mobile nodes upon entering the Wi-Fi network at 40km/h, between the prediction model (blue) and aggregate model (red); first period of arrival of mobile nodes
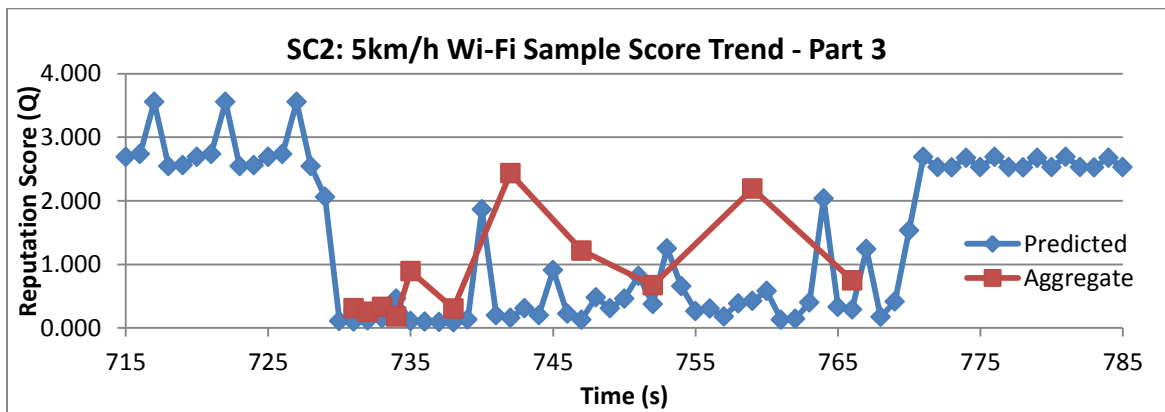

Figure 6.64: Scenario 2 comparison of the reputation scores sampled by the mobile nodes while within the Wi-Fi network at 40km/h, between the prediction model (blue) and aggregate model (red); first period of arrvial of mobile nodes
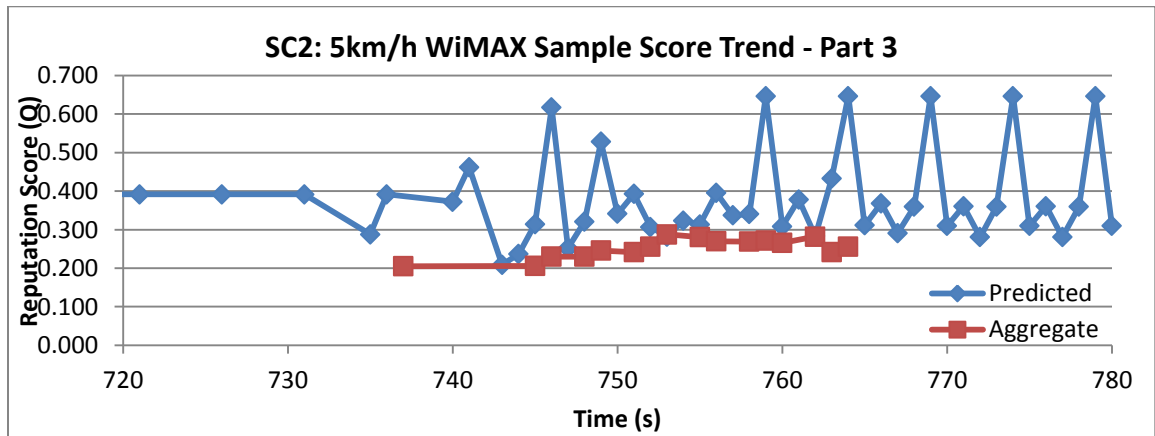
Figure 6.65: Scenario 2 comparison of the reputation scores sampled by the mobile nodes while within the WiMAX network at 40km/h, between the two models; during this period the first wave of mobile nodes are presented with the option to handover to Wi-Fi

During the second period of nodes joining the Wi-Fi network, between 108 and 120 seconds, the aggregate model facilitates 4 more handovers than the prediction model (15 vs. 11), as seen in Figure 6.62. This is due to the fact that as the nodes join and cause congestion, they drive the prediction model network scores down, as can be noted in Figure 6.67, where after 111s the majority of the scores calculated in Wi-Fi are below 0.5. Since the WiMAX scores are higher during this timeframe (Figure 6.68) for the prediction model, and significantly more nodes in the prediction model originally made the handover to WiMAX as compared to the aggregate model (refer to section 6.2.2.2), more nodes in the prediction model opt to stay in WiMAX. Conversely, since many mobile nodes in the aggregate model did not join the WiMAX network, they remained in the UMTS network with a fixed score of 0.08. As a result, many nodes in the aggregate model chose to join the Wi-Fi network.

The peaks seen in the sample score trend in Figure 6.66 are due to some nodes, upon entry into the Wi-Fi network, obtaining higher QoS; as seen here and in previous cases, this dampens over time due to resource contention. In the aggregate model these spike cases correspond to the peaks observed in the network advertised scores in Figure

6.67. In the prediction model, the initial dip at 110s in Figure 6.67 seems contrary to the sample scores calculated at 109s in Figure 6.66. This is due to the fact that the sample score values shown in the graph are the average of scores calculated at a given second. In this case, several scores were calculated approximately at 109s (with a variance of tens of milliseconds), and in sequence they were decreasing significantly. As a result, although the average value was plotted as a peak, the resulting prediction calculated at 110s was a dip. Subsequently, a similar but opposite behavior is seen at 110s in the sample scores, which leads to the prediction score returning to the pre-dip value. Following this anomaly, the network scores in the prediction model begin to align to the sample scores as they decrease and then stabilize around 116s.

Figure 6.66: Scenario 2 comparison of the reputation scores sampled by the mobile nodes while within the Wi-Fi network at 40km/h, between the prediction model (blue) and aggregate model (red); second period of arrival of mobile nodes
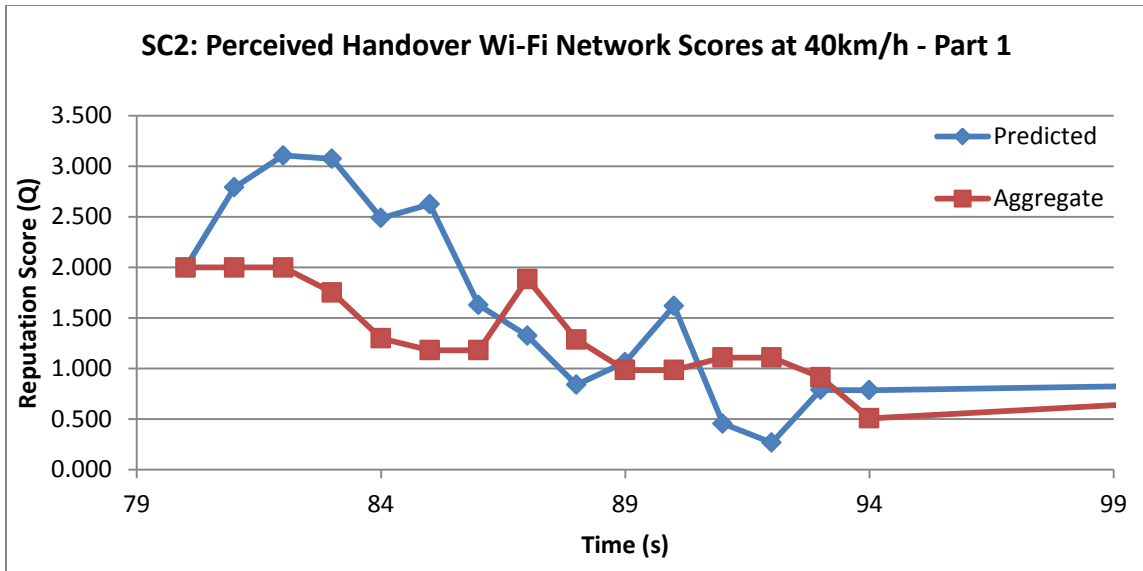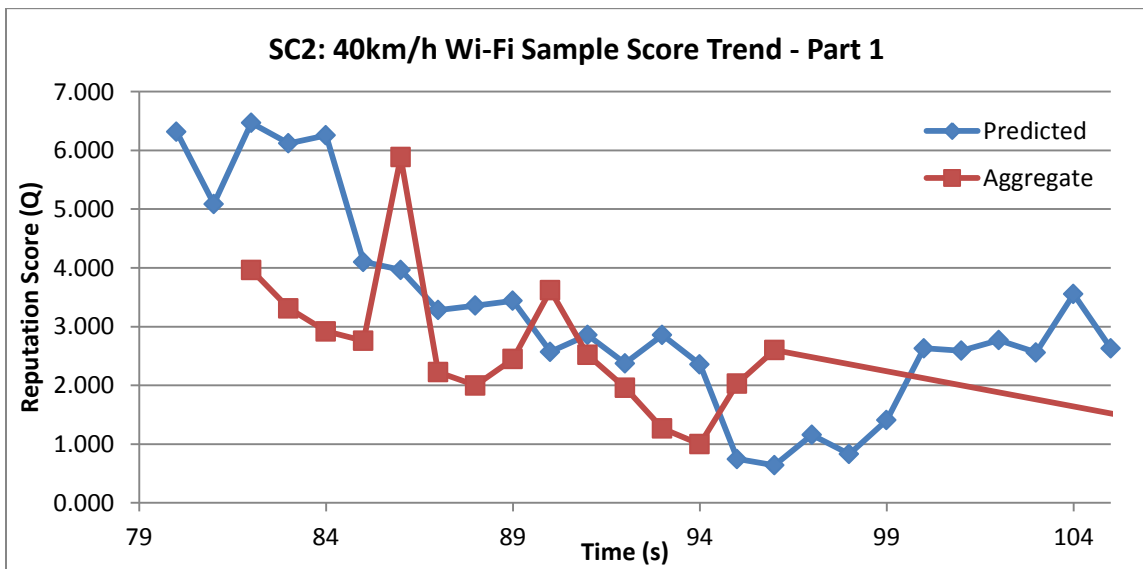
Figure 6.67: Scenario 2 comparison of the reputation scores perceived by the mobile nodes upon entering the Wi-Fi network at 40km/h, between the prediction model (blue) and aggregate model (red); second period of arrival of mobile nodes
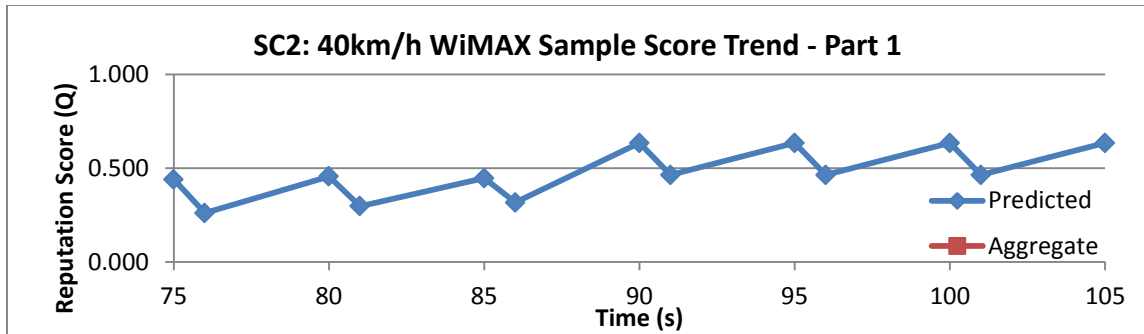

Figure 6.68: Scenario 2 comparison of the reputation scores sampled by the mobile nodes while within the WiMAX network at 40km/h, between the two models; during this period the second wave of mobile nodes are presented with the option to handover to Wi-Fi

The same can be seen during the third period between 130 and 141 seconds, where the initial congestion drives the predicted score down sharply (Figure 6.70), resulting in more nodes preferring to stay in WiMAX (scores seen in Figure 6.71) than making the switch to Wi-Fi, with the aggregate model facilitating 15 handovers, and the prediction model 12. Note that the peaks seen in the advertised scores (Figure 6.69) and

the sample scores are due to several nodes receiving good QoS upon entry, and then seeing this service diminish over time as more nodes join.

Figure 6.69: Scenario 2 comparison of the reputation scores perceived by the mobile nodes upon entering the Wi-Fi network at 40km/h, between the prediction model (blue) and aggregate model (red); third period of arrival of mobile nodes
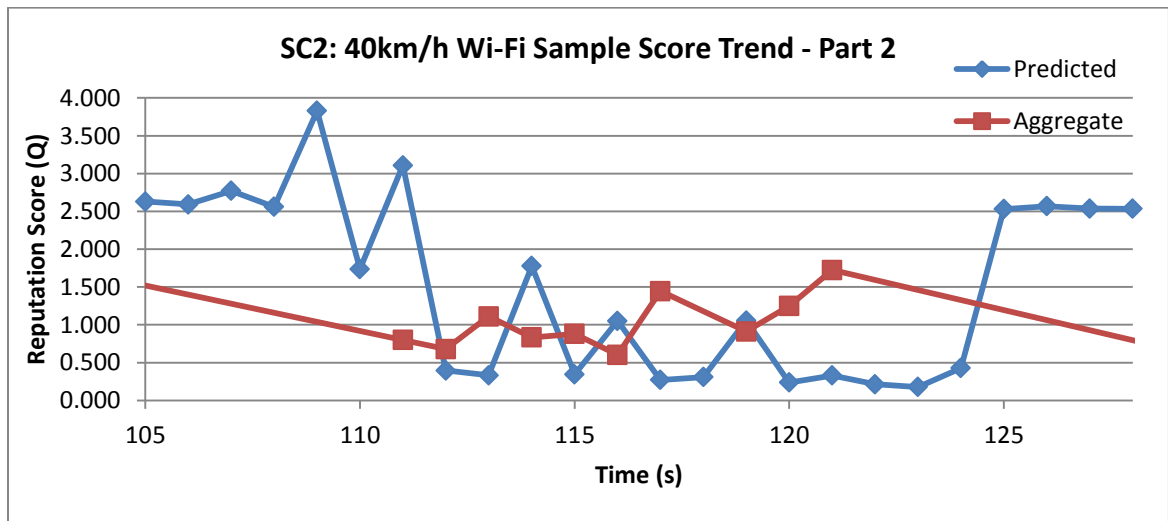

Figure 6.70: Scenario 2 comparison of the reputation scores sampled by the mobile nodes while within the Wi-Fi network at 40km/h, between the prediction model (blue) and aggregate model (red); third period of arrvial of mobile nodes

Figure 6.71: Scenario 2 comparison of the reputation scores sampled by the mobile nodes while within the WiMAX network at 40km/h, between the two models; during this period the third wave of mobile nodes are presented with the option to handover to Wi-Fi

Finally, during the last period between 148 and 180 seconds, as the last group of nodes begins to arrive, two other groups sequentially leave the Wi-Fi zone. In this scenario, again the aggregate model produces more handovers, 16, vs. 13 in the prediction model, as seen in Figure 6.62. In this case, as more nodes begin to leave the initial score calculated by the aggregate model tends to improve, but in the prediction model this initial score is balanced by the fact that existing nodes are still experiencing the congested network, as seen in Figure 6.73. In fact, since the network is improving gradually, the initial good scores calculated by a node upon entry tend to decrease immediately or soon after to the average. As a result, the prediction model advertises more conservative network scores due to the overall experience of all nodes at that moment, as can be observed in Figure 6.72.
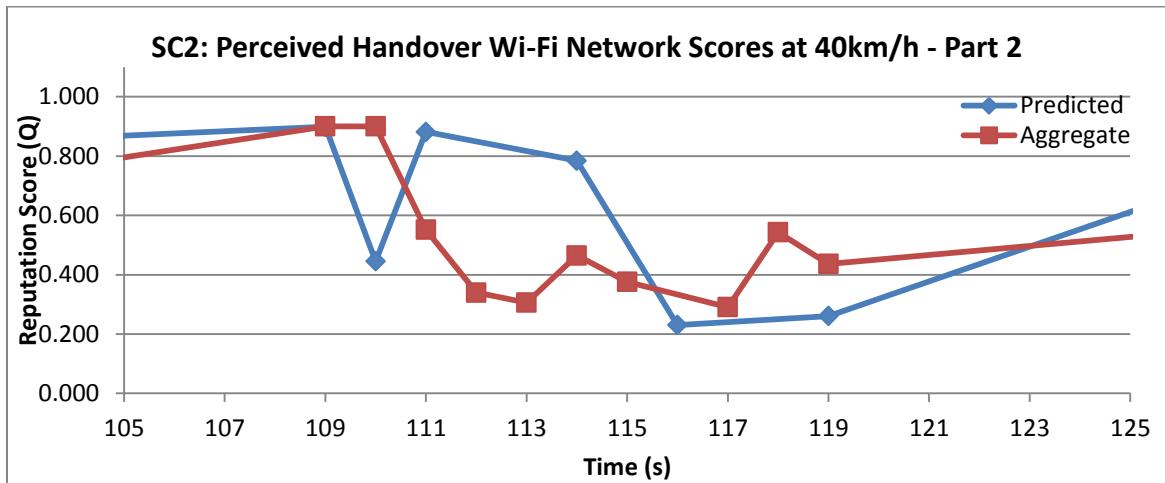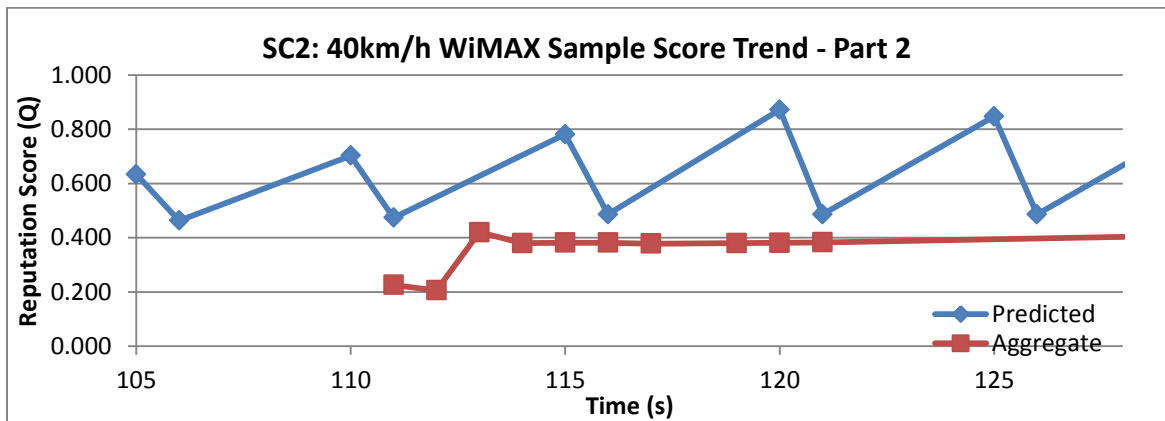
Figure 6.72: Scenario 2 comparison of the reputation scores perceived by the mobile nodes upon entering the Wi-Fi network at 40km/h, between the prediction model (blue) and aggregate model (red); fourth period of arrival of mobile nodes
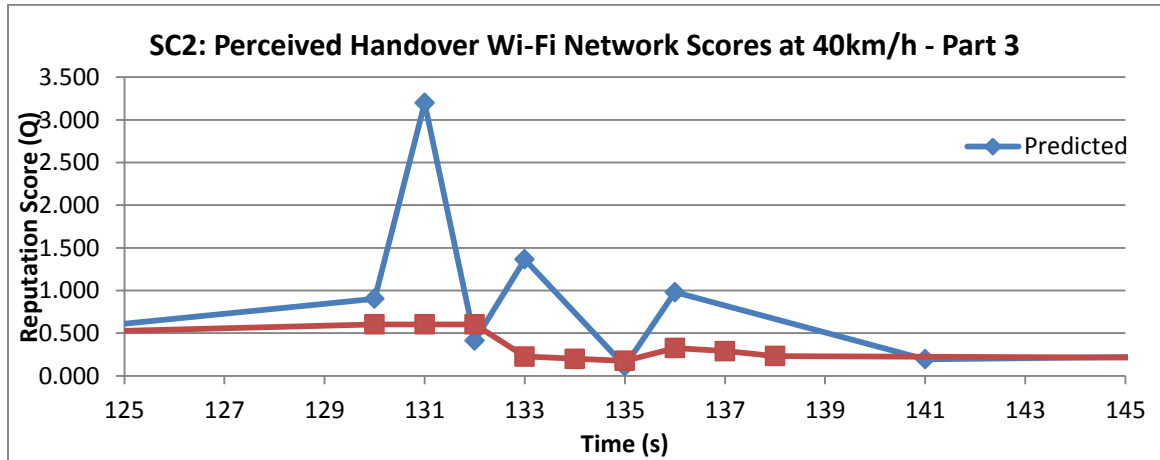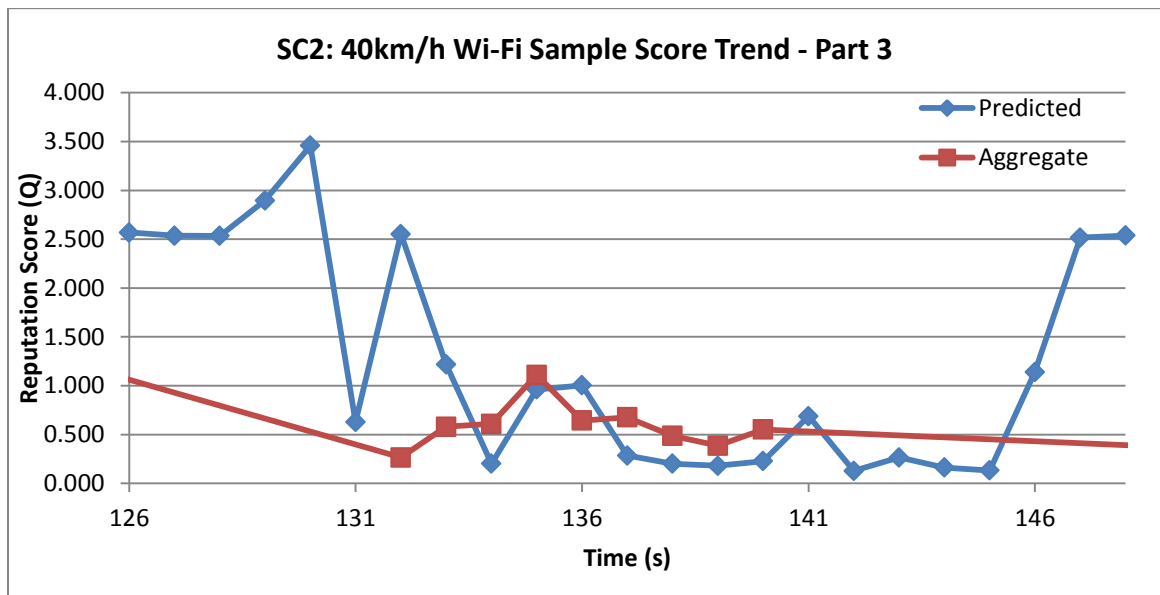

Figure 6.73: Scenario 2 comparison of the reputation scores sampled by the mobile nodes while within the Wi-Fi network at 40km/h, between the prediction model (blue) and aggregate model (red); fourth period of arrvial of mobile nodes

Figure 6.74: Scenario 2 comparison of the reputation scores sampled by the mobile nodes while within the WiMAX network at 40km/h, between the two models; during this period the fourth wave of mobile nodes are presented with the option to handover to Wi-Fi

6.3.3.3   Mobile Speed 75km/h:

At 75km/h, the pattern observed in the previous two speeds is consistent, in that there are more handovers facilitated through the aggregate model in comparison with the prediction model, as seen in Figure 6.75.
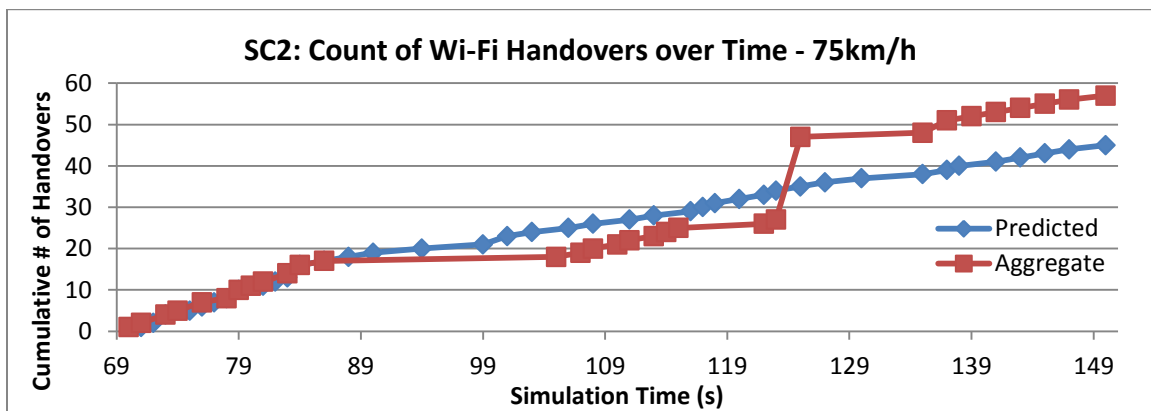
Figure 6.75: Scenario 2 comparison of the cumulative # of handovers over time between the proposed prediction model (blue) and the existing aggregate model (red) at 75km/h in the Wi-Fi network

At this speed, the nodes arrive at the Wi-Fi zone in roughly 2 distinct time periods. In the first period of arrivals, between 70 and 104s, there are actually more

handovers facilitated via the prediction model (24) vs. the aggregate model (17). The reason for this is that while both models are experiencing a declining network advertised score with the gradual joining of mobile nodes to the network (Figure 6.76), in the aggregate model two mobile nodes lose their connection to the Wi-Fi network soon after entry and as a result join the WiMAX network, as seen in Figure 6.78. In doing so, they sample very good scores, which is subsequently advertised by the WiMAX network agent. As a result, mobile nodes reaching the Wi-Fi network subsequent to this, which are in WiMAX, opt to stay in the network as the score is much higher (Wi-Fi sample scores can be seen in Figure 6.77). Through the course of this analysis though it has become evident that the initially calculated entry sample scores tend to be higher, so these aggregate values polled in the WiMAX network could be misleading as to the actual longer term service experienced by the mobile node.

Figure 6.76: Scenario 2 comparison of the reputation scores perceived by the mobile nodes upon entering the Wi-Fi network at 75km/h, between the prediction model (blue) and aggregate model (red); first period of arrival of mobile nodes

Figure 6.77: Scenario 2 comparison of the reputation scores sampled by the mobile nodes while within the Wi-Fi network at 75km/h, between the prediction model (blue) and aggregate model (red); first period of arrival of mobile nodes
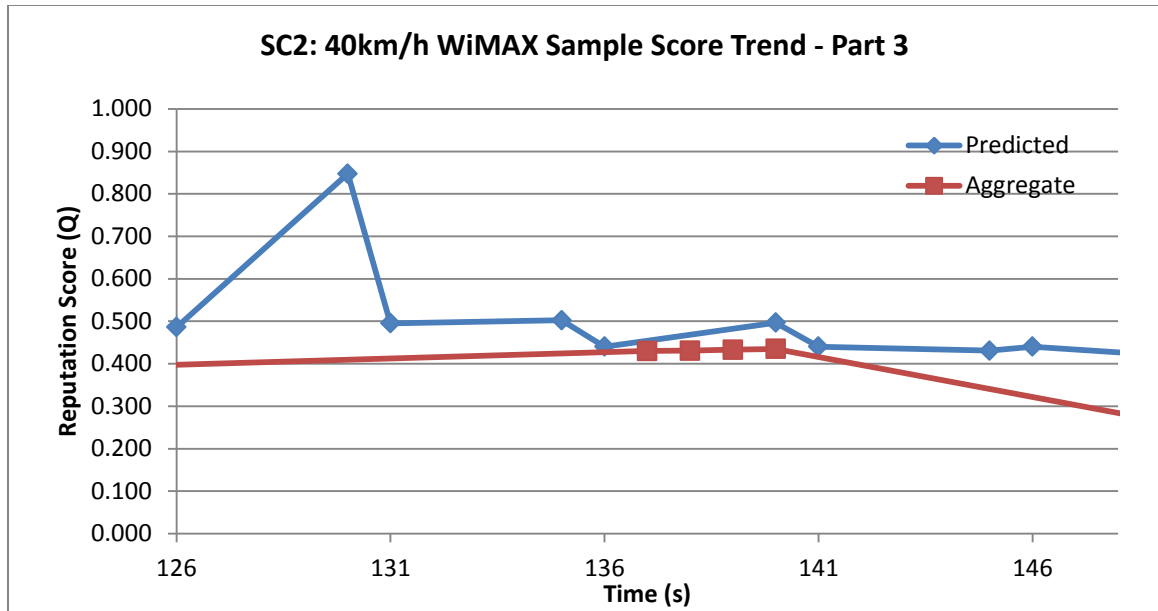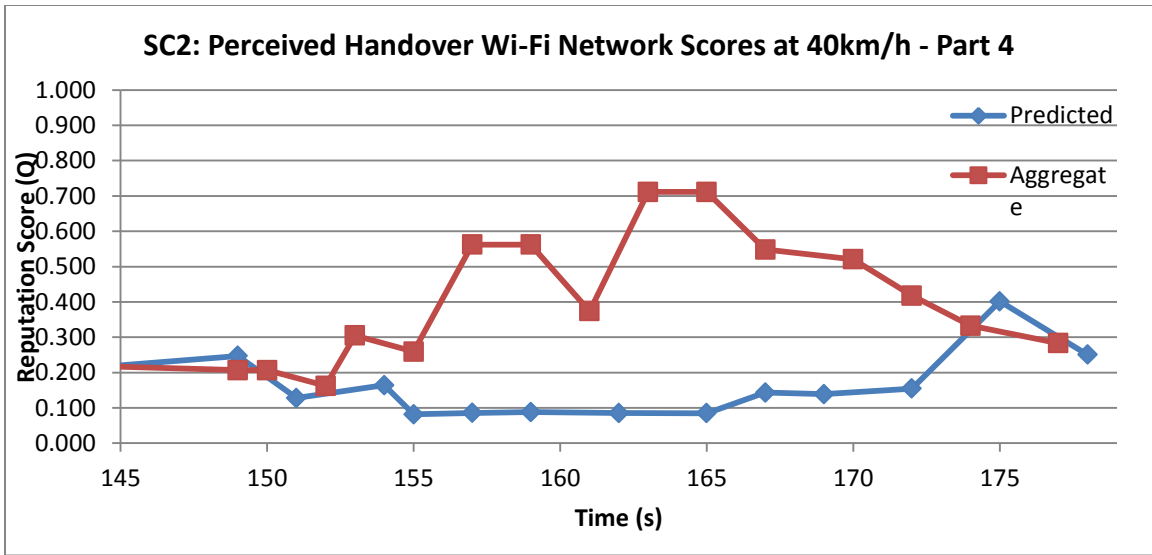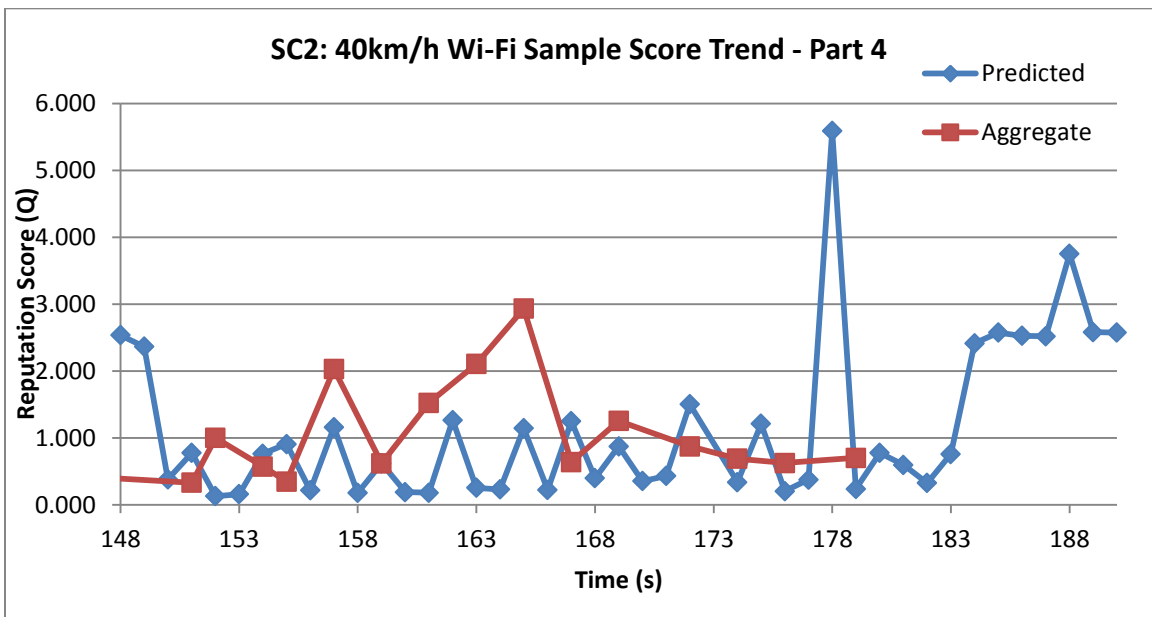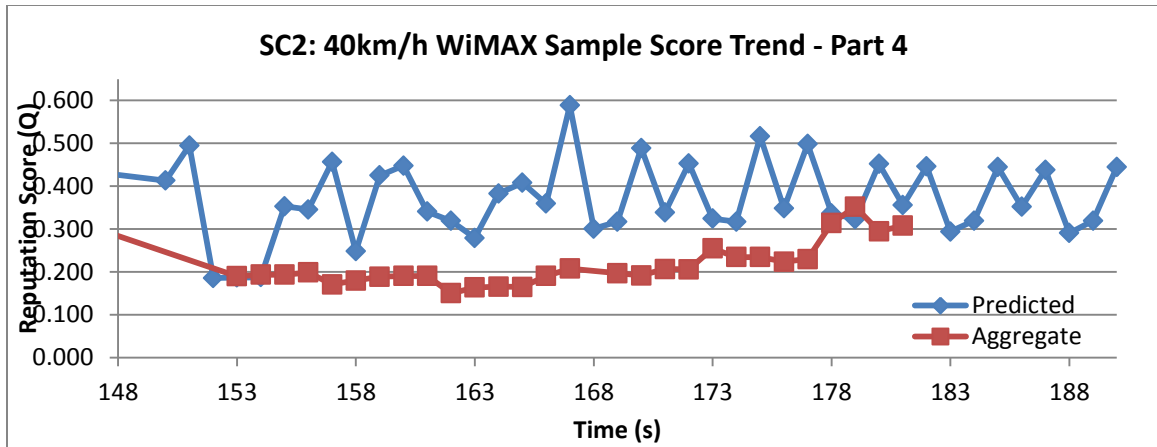

Figure 6.78: Scenario 2 comparison of the reputation scores sampled by the mobile nodes while within the WiMAX network at 75km/h, between the two models; during this period the first wave of mobile nodes are presented with the option to handover to Wi-Fi

In the second period, there are significantly more handovers to the Wi-Fi network in the aggregate mode, 40, vs. the prediction model, 21. The reason why any mobile nodes even join the Wi-Fi network is due to the fact that there were still many nodes which had not handed over to WiMAX initially and were in UMTS, with a network score of 0.08. As a result, when any such node reaches the Wi-Fi network, it opts to switch. Meanwhile, at 123s the first node exits the Wi-Fi coverage area and joins the WiMAX network, sampling a much lower score than the other two mobile nodes which had lost

their connection to Wi-Fi earlier in the first period, as seen in Figure 6.81. This brings the WiMAX network advertised score down significantly, and as a result both nodes in the WiMAX and UMTS coverage area in the aggregate model opt to go with the Wi-Fi network upon reaching its coverage zone.

Consequently, the aggregate scores beginning around 123s are steadily increasing due to the other two groups leaving the Wi-Fi coverage zone, and in general are higher than the prediction model, as observed in Figure 6.80. The higher scores are also due in part to the initially higher sample score perceived upon joining the network. In the prediction model, these initial high scores are dampened by the existing mobile nodes in the network which are still experiencing congestion, as can be seen in the network advertised scores in Figure 6.79.
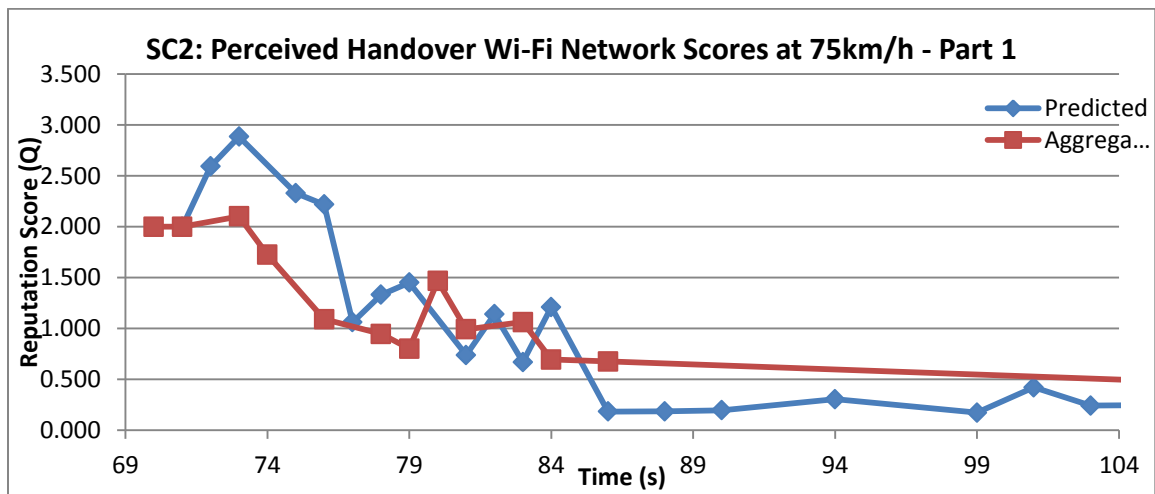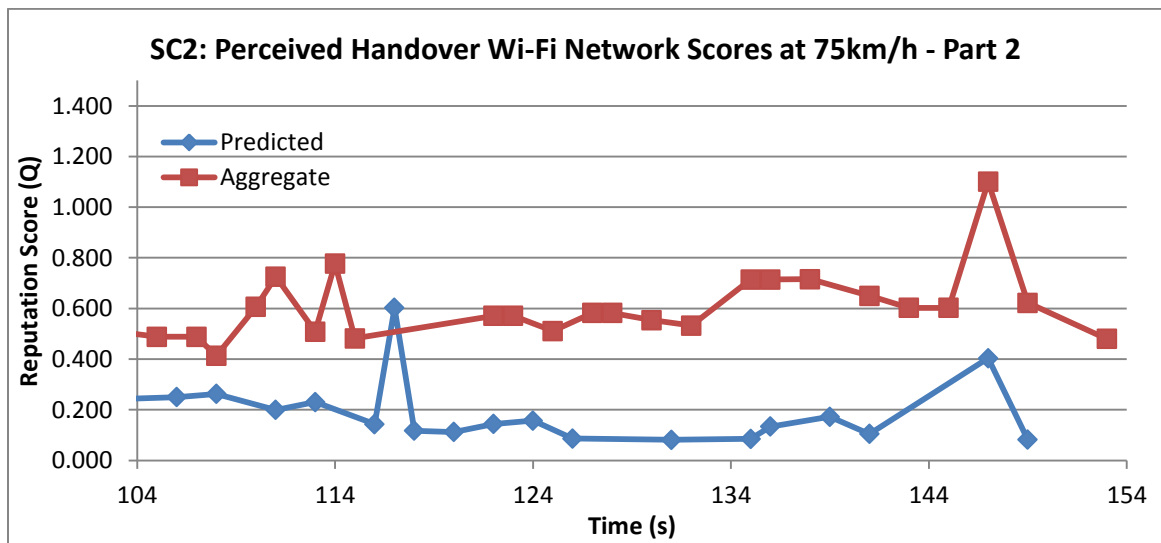

Figure 6.79: Scenario 2 comparison of the reputation scores perceived by the mobile nodes upon entering the Wi-Fi network at 75km/h, between the prediction model (blue) and aggregate model (red); second period of arrival of mobile nodes
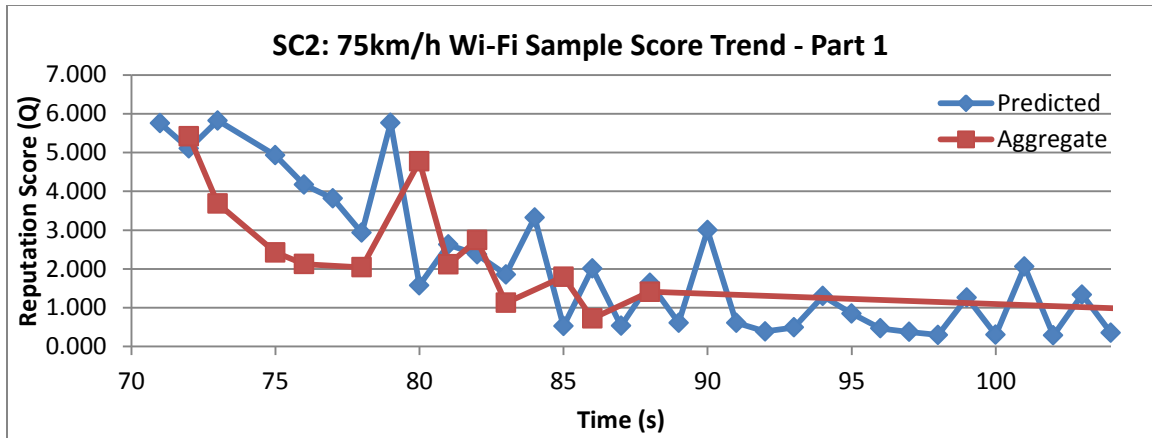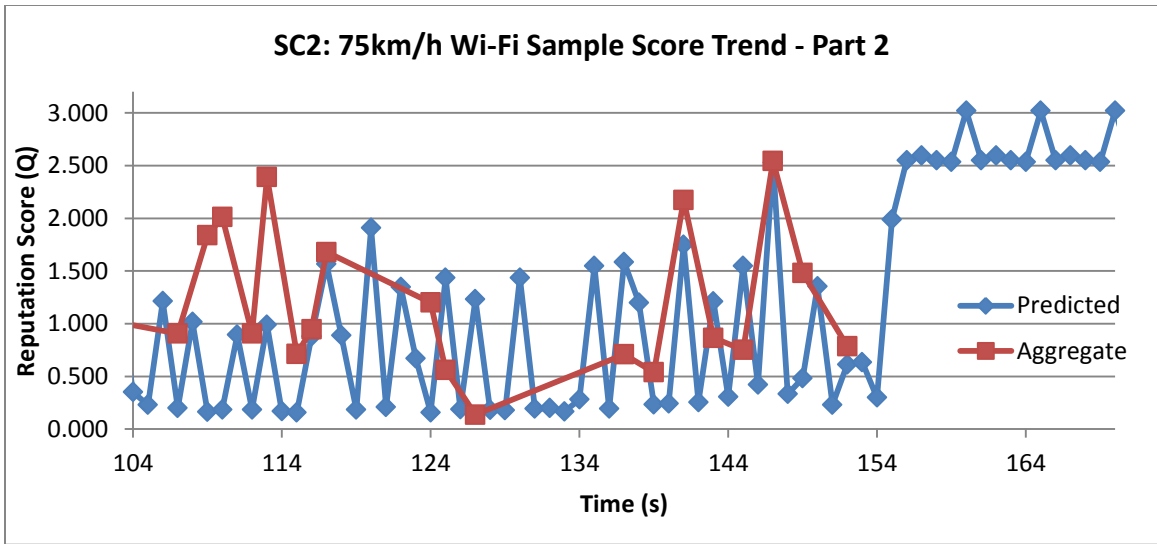
Figure 6.80: Scenario 2 comparison of the reputation scores sampled by the mobile nodes while within the Wi-Fi network at 75km/h, between the prediction model (blue) and aggregate model (red); second period of arrival of mobile nodes

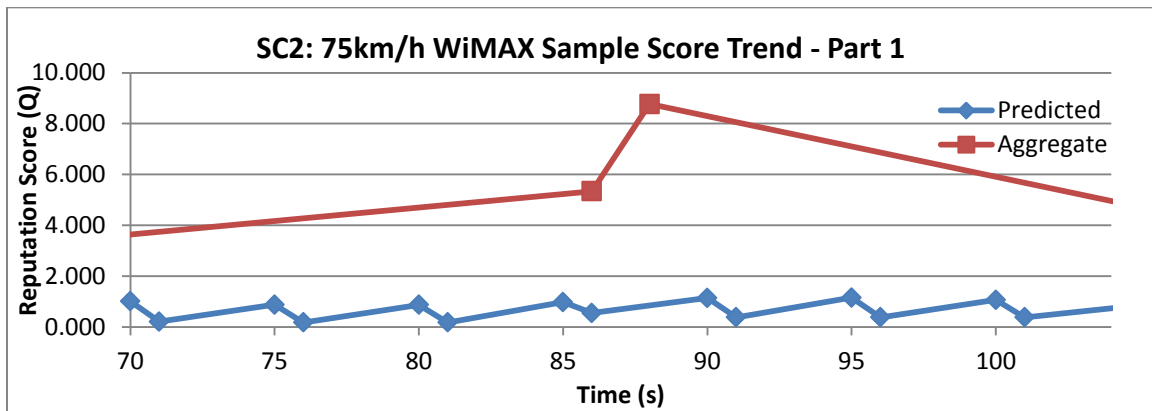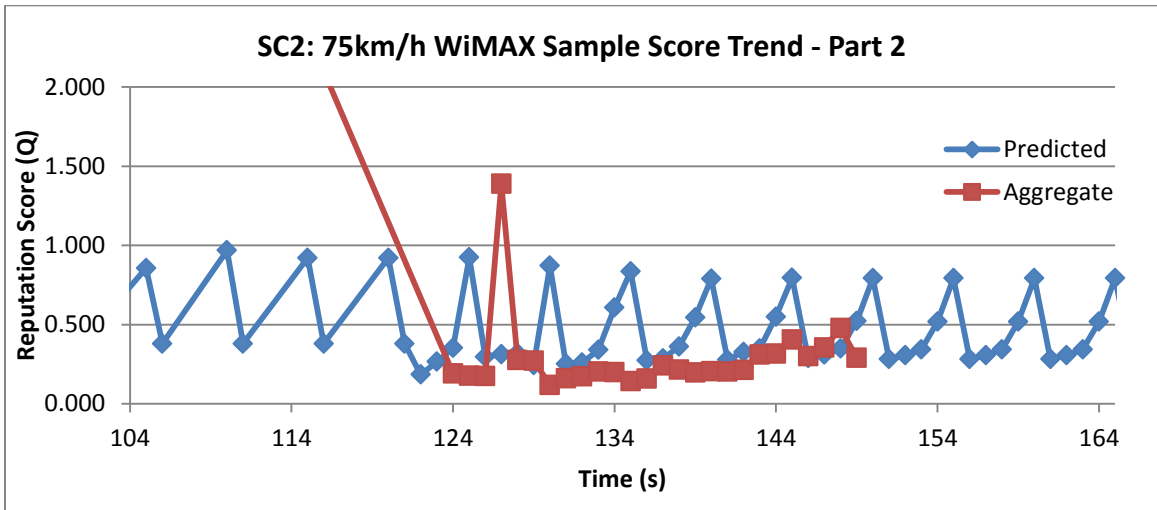**Figure 6.81: Scenario 2 comparison of the reputation scores sampled by the mobile nodes while within the WiMAX network at 75km/h, between the two models; during this period the second wave of mobile nodes are presented with the option to handover to Wi-Fi**

6.4   Summary

In this chapter the performance of the proposed model was analyzed through investigating the behavior of a single mobile node and that of multiple nodes in a

heterogeneous network environment. In both cases, the proposed model was compared against the original aggregation reputation score algorithm defined in [8], and varying traffic conditions were applied.

From a single node perspective it was observed that in a gradually varying traffic load, the proposed model provides advanced warning of a deteriorating network, allowing the mobile node to avoid handing over to a network that is on the cusp of performing below expectations. When traffic spikes are experienced the proposed model results in mobile nodes perceiving the network as providing acceptable QoS only after several good scores are experienced by mobile nodes in the network. This helps avoid VHOs during the initial recovery period after a traffic spike when there could still be volatility in the network.

In the multi-node scenarios, when a network is overloaded with traffic the proposed model will reduce the number of VHOs into the overloaded network, helping to balance the heterogeneous network environment. The sharper this traffic increase is, the more dramatic the reduction of handovers will be. Similarly to the single node scenario, when the traffic is reduced a given the network, this will not automatically correspond to more VHOs permitted by the network in question since the prediction model will in fact require multiple increasing values in order to report this in the network advertised scores. As a result, it provides a more conservative view on increasing network conditions.

7 Conclusion and Future Work

7.1 Conclusion

In synopsis, the proposed model improves upon the novel reputation based vertical handover scheme devised in [8] through the application of the GM(1,1) algorithm. This is accomplished through obtaining the raw reputation scores obtained from the mobile nodes traversing in a given network, calculating an aggregate score value, and then treating the sequence of multiple calculated aggregate values as the input into the GM(1,1) model. In doing so, we can see several benefits for an individual node and in the overall health of the networks leveraging this protocol in a heterogeneous environment.

For a single node, when there is gradually varying traffic load, the mobile node utilizing the prediction model scheme will have advanced notice of service degradation in networks discovered through the 802.21 protocol. This enables the mobile node to avoid making vertical handovers to networks which currently may have acceptable service but which will shortly offer unacceptable service. Furthermore, after a network surge in traffic occurs, the mobile node will not immediately consider joining the network. Rather, it will require several indications that the network is in an acceptable state before perceiving the network as fit to join.

In the case of a heterogeneous network environment with numerous nodes, when traffic is overloaded on a network, the prediction model will reduce the amount of VHO into the congested network, assuming that the current serving network is still performing well. This in turn improves the network conditions in the congested network and

balances the load throughout the heterogeneous ecosystem through taking into account the network advertised scores of the serving network and the target network. When the traffic is gradually increasing, the prediction model will have a much diminished impact on the overall number of handovers so long as the target network is providing a sufficient quality of service.

Furthermore, when a network's traffic is reduced, this will not immediately translate into more handovers facilitated by the prediction model. The important aspect to note is that the prediction model is always taking into consideration the network advertised scores of the current serving network and target network based on current sampled experiences by the mobile nodes utilizing the services. As a result, if a target network's traffic begins lessening, it will require a sufficient amount of mobile nodes within that network to sense the improvement in service before the advertised score reflects this. In the aggregate model, the advertised scores only reflect the scores calculated by the last node to join the network or leave through forced handover, so it does not necessarily paint the most realistic picture of the network. As a result, the prediction model behavior is seen as beneficial from a network management perspective since the overall score advertised by the network reflects the experience realized by all the nodes in the network as close to the point of request as possible.

Based on these results, it can be concluded that the prediction vertical handover algorithm enables an optimized selection of the appropriate network to maintain a high user experience, as well as to ensure that the user experience seen by all mobile nodes in heterogeneous ecosystem is well balanced through the use of up to date network advertised scores in both the serving and target networks.

7.2 Benefits

Given that a heterogeneous network ecosystem could implement the proposed solution outlined in this thesis, the benefit that would be yielded by each mobile node within the environment, in comparison with the original model, is an improved vertical handover decision making capability, i.e. the ability to make a better decision when faced with options. Furthermore, the enhanced algorithm would not impact the delay incurred by the decision making process since the complexity introduced by the proposed model is executed outside of the decision making process itself.

As a result, when the consumer is traversing through an area of overlapped heterogeneous network types, the mobile device will be able to make an improved decision on which network to join. This in turn will improve the overall experience perceived by the end user utilizing latency sensitive applications on the device, and also help to balance the load across the various network types in the ecosystem.

As the availability of open and free network access points grows within large metropolitan cities, the benefits realized by this approach will be more readily apparent, since the number of applicable overlapped heterogeneous networks with increase.

7.3 Future Work

In observing the behavior of the proposed model in comparison to the original reputation model under varying conditions, several ideas and improvements have already been devised that may help form the basis of future work to be carried out in this research area. The points below summarize these suggestions.

- **Expanded Comparison Set**: In addition to comparing the proposed model to the original aggregate model, a comparison can be made against other popular types of VHO methods such as MADM, Neural Network, Fuzzy Logic, and RSS to name several, which could further highlight the benefits of the model.

- **Investigate GM(1,1) Variances**: There are several more complex implementations of the GM(1,1) model such as Grey Residual Error Correction models using Fourier Transforms, back-propagating neural networks, and model algorithm control. It would be interesting to see if any of these yield significant performance benefits over the simple GM(1,1) model used.

- **Improve Score Noise**: In analyzing the sample score and perceived network score data sets, it is evident that there is a significant amount of oscillatory behavior due to numerous factors. An area for improvement here is to see if there is a way to apply feedback in the score or prediction calculation process in order to smooth out the oscillation.

- **Expand Traffic Analysis**: The experimentation carried out to date has covered model comparisons leveraging conversational CoS traffic sources. Understanding the performance of the proposed model under different CoS traffic sources, or in environments where there are several types of CoS traffic sources being concurrently used would be beneficial.

8   References

1.   *IEEE Standard for Local and Metropolitan Area Networks- Part 21: Media Independent Handover.* IEEE Std 802.21-2008, 2009, p. 1-301.

2.   Cisco Systems, I., *Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2011–2016*, 2012. p. 29.

3.   *Wi-Fi Alliance Organization.* 2012 [cited 2012 May 1, 2012]; Brief history of the Wi-Fi Alliance organization and several facts about Wi-Fi]. Available from: http://www.wi-fi.org/about/organization.

4.   Andrews, J.G., A. Ghosh, and R. Muhamed, *Fundamentals of WiMAX : understanding broadband wireless networking.* Prentice Hall communications engineering and emerging technologies series2007, Upper Saddle River, NJ: Prentice Hall. xxvii, p. 1-449 .

5.   Bo, L., et al., *A Survey on Mobile WiMAX [Wireless Broadband Access].* Communications Magazine, IEEE, 2007. **45**(12): p. 70-75.

6.   Pareit, D., et al., *The History of WiMAX: A Complete Survey of the Evolution in Certification and Standardization for IEEE 802.16 and WiMAX.* Communications Surveys & Tutorials, IEEE, 2011, p. 1-29.

7.   Molisch, A.F., *Wireless communications.* 2nd ed. ed2011, Oxford: John Wiley & Sons.

8.   Zekri, M., B. Jouaber, and D. Zeghlache. *On the use of network QoS reputation for vertical handover decision making.* in *GLOBECOM Workshops on Handover and Network Selection in Future Broadband Networks, 2010 IEEE.*

9.   Taniuchi, K., et al., *IEEE 802.21: Media independent handover: Features, applicability, and realization.* Communications Magazine, IEEE, 2009. **47**(1): p. 112-120.

10.   De La Oliva, A., et al., *An overview of IEEE 802.21: media-independent handover services.* Wireless Communications, IEEE, 2008. **15**(4): p. 96-103.

11.   Nyamapfene, A. and B. Naeem, *Seamless Vertical Handover in Wi-Fi and WiMAX Networks using RSS and Motion Detection: An Investigation.* The Pacific Journal of Science and Technology, 2011. **12**(1): p. 298 – 304.

12. Bathich, A.A., M.D. Baba, and M. Ibrahim. *IEEE 802.21 based vertical handover in WiFi and WiMAX networks*. in *Computers & Informatics (ISCI), 2012 IEEE Symposium on*. 2012.

13. Sueng Jae, B., C. Min Young, and S. Jungmin. *Handover triggering mechanism based on IEEE 802.21 in heterogeneous networks with LTE and WLAN*. in *Information Networking (ICOIN), 2011 International Conference on*. 2011.

14. Mohamed, L., C. Leghris, and A. Adib. *A Hybrid Approach for Network Selection in Heterogeneous Multi-Access Environments*. in *New Technologies, Mobility and Security (NTMS), 2011 4th IFIP International Conference on*. 2011.

15. Yussuf, A.A., W.H. Hassan, and S. Issa. *A review of VHD approaches in next generation wireless networks*. in *Digital Information and Communication Technology and it's Applications (DICTAP), 2012 Second International Conference on*. 2012.

16. Alkhawlani, M.M., K.A. Alsalem, and A.A. Hussein. *Multi-criteria Vertical Handover by TOPSIS and fuzzy logic*. in *Communications and Information Technology (ICCIT), 2011 International Conference on*. 2011.

17. Rakovic, V. and L. Gavrilovska. *Novel RAT selection mechanism based on Hopfield neural networks*. in *Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT), 2010 International Congress on*. 2010.

18. Deng, J.L., *Introduction to Grey system theory*. J. Grey Syst., 1989. **1**(1): p. 1-24.

19. Kayacan, E., B. Ulutas, and O. Kaynak, *Grey system theory-based models in time series prediction*. Expert Systems with Applications, 2010. **37**(2): p. 1784-1789.

20. Hong, W. and C. Fuzhong. *Chinese Exchange Rate Forecasting Based on the Application of Grey System DGM(2, 1) Model in Post-Crisis Era*. in *Information Management, Innovation Management and Industrial Engineering (ICIII), 2010 International Conference on*. 2010.

21. Wuji, S., L. Yichen, and F. Wei. *Predictions based on the GM model of grey system*. in *Information Theory and Information Security (ICITIS), 2010 IEEE International Conference on*. 2010.

22.   Zhengxin, W., D. Yaoguo, and W. Yemei. *A new grey Verhulst model and its application*. in *Grey Systems and Intelligent Services, 2007. GSIS 2007. IEEE International Conference on*. 2007.

23.   Jin, S. and Z. Jianyong. *The Use of Grey Verhulst Model in the Prediction of Operating Activity Cash Flow*. in *Business Intelligence and Financial Engineering (BIFE), 2011 Fourth International Conference on*. 2011.

24.   Jichen, F., et al. *Forecasting model of automobile logistics demand based on Gray Residual-BP neural network*. in *Image and Signal Processing (CISP), 2010 3rd International Congress on*. 2010.

25.   Peng, G. and L. Tengfei. *Multivariable MAC control based on residual error gray feedback correction and its application*. in *Automation and Logistics, 2008. ICAL 2008. IEEE International Conference on*. 2008.

26.   Fei, Z.-h., Z.-y. Jiang, and Z.-f. He. *Discrete GM(1,1) model and its application for forecasting of real estate prices*. in *Management Science and Industrial Engineering (MSIE), 2011 International Conference on*. 2011.

27.   Chiung-Wen, H. and C. Pao-Long. *Using grey model to predict the achievement of policy objectives*. in *Technology Management in the Energy Smart World (PICMET), 2011 Proceedings of PICMET '11:*. 2011.

28.   Li, J. *Application of improved grey GM(1,1) model in tourism revenues prediction*. in *Computer Science and Network Technology (ICCSNT), 2011 International Conference on*. 2011.

29.   Yan, B., M. Ke, and R. Qingchang. *A NN-GM (1,1) model - based analysis of network traffic forecasting*. in *Broadband Network and Multimedia Technology (IC-BNMT), 2010 3rd IEEE International Conference on*. 2010.

30.   Yuan, L. and C. Jianhua. *Network Traffic Prediction Based on Grey Neural Network Integrated Model*. in *Wireless Communications, Networking and Mobile Computing, 2008. WiCOM '08. 4th International Conference on*. 2008.

31.   *NS-2 Homepage*. 2011 [cited 2010 September 1, 2010]; The main web site for the NS-2 simulator, with version history, help guides, and descriptions of contributed code. Available from: http://nsnam.isi.edu/nsnam/index.php/Main_Page.

32.   *NIST Seamless and Secure Mobility Homepage*. 2012 [cited 2010 September 1, 2010]; The main page for downloading the NIST mobility module for NS-2, and for reviewing FAQs and help guides. Available from: http://www.nist.gov/itl/antd/emntg/ssm_tools.cfm.

33.   *EURANE Homepage*. 2006 [cited 2010 September 1, 2010]; Homepage for the EURANE contributed code for NS-2, contains access to help guides and FAQs. Available from: http://eurane.ti-wmc.nl/eurane/.

34.   Matrawy, A. and I. Lambadaris. *MPEG4 Traffic Source Homepage*. 2006 [cited 2010 September 1, 2010]; Home page for downloading the source code for the NS-2 MPEG4 traffic generator. Available from: http://www.sce.carleton.ca/~amatrawy/mpeg4/.

35.   Stevens-Navarro, E. and V.W.S. Wong. *Comparison between Vertical Handoff Decision Algorithms for Heterogeneous Wireless Networks*. in *Vehicular Technology Conference, 2006. VTC 2006-Spring. IEEE 63rd*. 2006.

36.   Cisco Systems, I., *Cisco TelePresence Network Systems 2.0 Design Guide*2009. 276.

37.   Greengrass, J., J. Evans, and A.C. Begen, *Not All Packets Are Equal, Part 2: The Impact of Network Packet Loss on Video Quality*. Internet Computing, IEEE, 2009. **13**(2): p. 74-82.