

An Economic Model of Subscriber Offloading Between Mobile Network Operators and a WLAN Operator

Cameron W. Patterson

Thesis submitted to the Faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of

Master of Science

in

Electrical Engineering

Allen B. MacKenzie, Chair

Luiz A. DaSilva

R. Michael Buehrer

September 17, 2014

Blacksburg, Virginia

Keywords: Cellular, WLAN, Wi-Fi, heterogeneous wireless network, pricing, economic
model, optimization

Copyright 2014, Cameron W. Patterson

An Economic Model of Subscriber Offloading Between Mobile Network Operators and a WLAN Operator

Cameron W. Patterson

(ABSTRACT)

With increasing mobile data demand there is a push towards heterogeneous networks. Small-scale operators (SSOs) of WLANs are becoming more prevalent, while Mobile Network Operators (MNOs) seek an outlet for their customers' data usage. These conditions prompt the need for an effective relationship between the two parties for the purpose of offloading cellular data traffic to WLANs in a way that is economically beneficial to all involved. This thesis presents a model of such a relationship, in which the SSO sets a strategic offloading price per subscriber and several MNOs can choose how many subscribers they want to offload in order to minimize their costs. We determine the optimal offloading price, identify how the SSO incorporates its own network's quality of service (QoS) into its price decision, and examine the way in which the MNOs' cost structures affect their ability to offload. This model can be applied by both MNOs and SSOs to make informed network deployment decisions, even before engaging in an offloading relationship.

Acknowledgements

I would like to express my appreciation for the members of my thesis committee: Dr. Allen B. MacKenzie, Dr. Luiz A. DaSilva, and Dr. R. Michael Buehrer. In particular, I would like to thank Dr. MacKenzie and Dr. DaSilva for providing guidance and support in all of my research efforts.

My gratitude also extends to all of my family and friends for the support they have shown me throughout this whole process. Among these friends are my labmates with whom I have enjoyed interacting both socially and professionally. In particular, I would like to acknowledge Varuni Sastry for aiding me in beginning my research.

Finally, I would like to thank the National Science Foundation, as this material is based upon work supported by the NSF under Grant No. 1147790. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the National Science Foundation.

Contents

List of Figures	viii
List of Tables	x
1 Introduction	1
1.1 Motivation for Research	1
1.2 Summary of Exploration	4
1.3 Research Contributions	5
2 Background and Related Work	7
2.1 Local Access Wireless Networks	8
2.1.1 Wi-Fi	8
2.1.2 Femtocells	10

2.2	History and Current State of Wi-Fi Offloading	11
2.3	Hotspot 2.0	14
2.4	Device Handoff	16
2.5	Delay-Tolerant Data and Opportunistic Offloading	19
2.6	Economic Models for Offloading	23
2.7	Viability of WLAN Offloading as a Long-term Solution	27
2.8	Conclusion	28
3	Offloading Model and Primary Exploration	30
3.1	Model and Problem Overview	31
3.1.1	Notation and Fundamental Relationships	31
3.1.2	MNO Cost Function	33
3.1.3	Behavior of an Individual MNO	34
3.1.4	Problem Outline	36
3.2	Piecewise-Quadratic Cost Function	38
3.2.1	Number of Users Offloaded	38
3.2.2	Simplified SSO Utility: Revenue Only	39
3.2.3	Full SSO Utility	42

3.3	Direct Implications of the Model	44
3.4	Generalized Inferences	48
3.4.1	MNO Characteristics That Affect Offloading	48
3.4.2	Balance of Offloading Among Several MNOs	49
3.4.3	SSO with Limited Information	55
3.4.4	Application	58
3.5	Conclusion	59
4	Different Cost Functions and Further Exploration	60
4.1	General Polynomial Form Cost Function	61
4.1.1	Number of Users Offloaded	62
4.1.2	Optimizing Prices	63
4.1.3	Comparison of Results With Quadratic Cost Function	64
4.2	Exponential Cost Function	65
4.2.1	Number of Users Offloaded	66
4.2.2	Optimizing Prices	67
4.2.3	Comparison of Results With Quadratic Cost Function	67
4.3	Case Study of Offloadable Users	69

4.4	Outline of a Time-Evolving Offloading Model	73
4.4.1	Detailed Changes to the Model	74
4.4.2	Preview of Analysis	75
4.5	Conclusion	77
5	Conclusion	78
	Bibliography	81
A	Derivations for Results in Sections 4.1 and 4.2	89
A.1	Revenue-optimizing Price for Polynomial Cost Function	89
A.2	χ -intercept for Polynomial Cost Function	90
A.3	Revenue-optimizing Price for Exponential Cost Function	91
A.4	χ -intercept for Exponential Cost Function	92

List of Figures

3.1	MNO cost function with a strictly convex increase after surpassing a threshold number of users.	34
3.2	The SSO's decision space for setting the offloading price χ is split up into $M + 1$ regions.	37
3.3	Difference between $\chi_{G_k}^*$ and $\chi_{k,G}^*$	41
3.4	Equilibrium number of MNO users offloaded from the cellular network to the SSO's WLAN, as the total number of active MNO users (n_{mno}) varies. n_G^* is the number of offloaded users when the SSO is acting on its revenue function alone, and n_U^* is the number when the SSO acts on its full utility function.	45
3.5	Equilibrium offloading price per user, set by the SSO, as the total number of active MNO users (n_{mno}) varies. χ_G^* is the price when the SSO is acting on its revenue alone, and χ_U^* is the price when the SSO acts on its full utility function.	45

3.6	Equilibrium number of MNO users offloaded n_U^* and price χ_U^* per offloaded user, as the number of native SSO customers (n_{ss0}) varies.	46
3.7	The globally optimal price χ_U^* falls in Region 1, where only a single MNO will offload.	51
3.8	The globally optimal price χ_U^* falls in Region 3.	51
3.9	The effect of the variation of MNO parameters on the offloading scenario. . .	52
3.10	Maximizing the SSO's revenue function dictates a price that entices two MNOs to offload, whereas the full utility function dictates a price that would entice only one MNO to offload.	53
3.11	Maximizing the SSO's revenue function dictates a price that entices three MNOs to offload, whereas the full utility function dictates a price that would entice only two MNOs to offload.	53
4.1	Area in downtown Oulu, Finland observed in this case study. We assessed the WLAN coverage within this area and determined the achievable offloading benefit when our model is applied. This figure is published in [1], ©2014 IEEE.	70
4.2	W_{min} , the minimum number of offloadable users, is greater than the number of users the MNO would offload for both the SSO's utility-maximizing case and revenue-maximizing case. This figure is published in [1], ©2014 IEEE. .	72

List of Tables

3.1	Single MNO Simulation Parameters	44
3.2	Limited information simulation results (less restrictive accuracy requirements).	56
3.3	Limited information simulation results (more restrictive accuracy requirements).	57
4.1	Case Study Parameters	69

Chapter 1

Introduction

1.1 Motivation for Research

The continued rapid growth of mobile broadband data traffic [2] places significant monetary burdens on Mobile Network Operators (MNOs). The MNOs are driven to invest in additional infrastructure in order to increase their network capacity, but they encounter two significant economic obstacles. The first obstacle is encountered when using smaller cell sites to take advantage of frequency reuse, a common strategy in densely populated areas. Though the standalone cost of an individual Micro Base Station (BS) is lower than that of a Macro BS,

Parts of this chapter are based on the “Introduction” section of [1], ©2014 IEEE.

the cost per unit area of coverage is greater if Micro BSs are used [3]. When attempting to increase spatial reuse, there is potentially increased cost per unit area of coverage.

The second and possibly more damaging economic obstacle to network upgrades is a decrease in MNO revenue per unit traffic. This effect is explored in [4], in which a Nordic MNO offers unlimited data plans and its average customer continues to consume more data at the same price. These challenges make future network upgrades more difficult, which could cause a decline in the average quality of service experienced by customers.

The narrower profit margin is especially problematic for some European MNOs, who, as a result of extremely costly licensing investments, would need a significantly higher average revenue per user (ARPU) just to get a reasonable return on their 3G networks [5] [6]. This has contributed to the same European operators lagging far behind American operators in 4G deployment despite their considerable head-start on 3G deployment [7].

In response to these difficulties, many MNOs (e.g. AT&T Wireless and Verizon Wireless in the United States) have removed their unlimited data plans in an effort to curb their customers' average data use. Another option, which is only beginning to be explored, is to make use of large-scale Wi-Fi deployment by offloading cellular traffic to wireless local area networks (WLANs) in dense environments. The use of WLAN infrastructure offers significantly lower deployment and operational costs than any type of cell site [4]. Widely-deployed WLANs in densely populated areas can decrease the burden on cellular networks while continuing to satisfy large data demands from mobile customers.

Concurrent with changes in the mobile data landscape, there have been a number of advancements in WLAN connectivity and heterogeneous networks that can be considered part of the aforementioned effort to relieve load on 3G and 4G cellular networks. Offloading large amounts of traffic onto WLANs can boost overall capacity due to increased spatial reuse in WLANs and potential access to additional, underutilized spectrum (such as 802.11a in the 5 GHz ISM band). This endeavor has been the object of much research, with varying opinions as to its viability. The authors of [8] are skeptical of significant offloading benefits due to the practical issue of WLAN deployments being optimized for laptop Wi-Fi devices, rather than the more constrained devices found in mobile phones. Other research, however, presents convincing arguments to the contrary. A dense urban WLAN deployment is studied in [9], which concludes that such a deployment would likely decrease network outage time and increase average user throughput as compared to femtocell deployment or the sole use of macrocells. There is also interest in the more technical aspects of WLAN offloading, such as the circumstances under which Wi-Fi can be used as a trusted replacement for cellular deployments [10] and methods of targeting specific types of traffic for offloading (such as delay-tolerant traffic in [11] and [12]).

We expect that these continued research efforts, combined with industry advances, will facilitate conditions that allow for easy transitions between cellular and Wi-Fi access. This scenario could benefit from an economic analysis of WLAN offloading. MNOs and WLAN operators would be more likely to engage in economic relationships if they are presented with justified strategies that are effective at securing their interests.

1.2 Summary of Exploration

This thesis presents a practical model for MNOs to offload some of their subscribers to WLANs. We refer to the WLAN operators discussed here as Small Scale Operators (SSOs) due to the relative size and coverage area of their networks in comparison to cellular networks. Examples of such SSOs include commercial providers of WLAN connectivity (e.g., Boingo) as well as university campuses and municipalities that have deployed extensive Wi-Fi infrastructure.

In the model presented in this thesis, the SSO sets an offloading price per user and each MNO makes a decision regarding how many of its own active subscribers to offload. In choosing a price, the SSO must weigh the revenue it will receive against the quality of service (QoS) strain that the additional users will place on its network. An MNO must simply compare the price to offload a user with the marginal cost to service this user natively on the cellular network.

A significant focus of this thesis is on finding the optimal offloading price that the SSO should charge in order to maximize its objective function. We begin by finding its revenue-maximizing price, but we ultimately focus on optimizing its full utility function, which includes a term that acts as a hedge against the QoS burden that offloaded users will place on the SSO's network. In both scenarios, the SSO must anticipate how the MNO will react before setting its price. The MNO uses a cost function to determine the marginal cost of servicing additional users on its network. The SSO considers the relationship between the

price and the number of users that the MNO(s) will offload and uses it to determine the optimal price.

One of the first results we examine is how the SSO's utility-optimizing strategy differs from its revenue-optimizing strategy. The QoS term in the SSO's utility function has a visible effect on how the SSO chooses its offloading price. In addition to this, we consider specifically how an MNO's cost function parameters influences its offloading tendencies, and how this behavior can affect the offloading options of other MNOs who have a relationship with the same SSO. We also explore how the different forms of the MNO's cost function can affect the offloading relationship. Finally, we offer an outline of a time-evolving model of the same offloading relationship as an early effort towards future work on this topic.

1.3 Research Contributions

The key contributions of this thesis are as follows:

- We present a new model of the relationship between an SSO and multiple MNOs in the context of WLAN offloading, identifying the incentives for each party and determining the utility maximizing offloading price that the SSO should charge.
- We derive conclusions on the nature of this relationship in a real-world offloading scenario with multiple MNOs. One MNO's cost structure and population of active

subscribers can affect both the SSO's strategy in setting the price and the ability of other MNOs to offload their subscribers.

- We offer a use case for the application of this offloading relationship in network planning. In anticipation of the availability and cost of WLAN offloading, an MNO can target a certain capacity limitation for a new network deployment so as to minimize the total cost of this network over its lifetime of operation.
- We present a case study on the practicality of our model in a real-world environment, using data from an already-deployed public Wi-Fi network and determining the proportion of users in a particular area that can be covered by Wi-Fi.

The work presented in this thesis can enable both MNOs and SSOs to make informed decisions regarding network deployment based on the projected cost/benefits of engaging in an offloading relationship.

Chapter 2

Background and Related Work

In order to fully understand the offloading model presented in this thesis we must first consider its context. In this chapter we give a summary of WLAN technology, including its technical evolution and its progress towards being used as a mobile supplement. We also provide an overview of academic work related to this area, including other economic models for offloading traffic.

Parts of this chapter are based on the “Background and Related Work” section of [1], ©2014 IEEE.

2.1 Local Access Wireless Networks

2.1.1 Wi-Fi

The term “Wi-Fi” is trademarked by the Wi-Fi Alliance. Wi-Fi devices are defined as “wireless local area network (WLAN) products that are based on the Institute of Electrical and Electronics Engineers (IEEE) 802.11 standards.” (The term “Wi-Fi” has become synonymous with “WLAN” and will be used as such in this thesis.) The IEEE 802.11 standards set defines the physical layer (PHY) and media access control (MAC) layer specifications for WLANs and has evolved significantly since its original release in 1997. The original standard allowed for a relatively low peak data rate (by today’s standards) of 2 megabits per second (Mbps) but has evolved to allow practical transfer rates of tens and sometimes hundreds of Mbps, depending on the version in use, the channel conditions, and the network settings.

The first PHY revisions to the standard were the ‘a’ and ‘b’ amendments in 1999. 802.11a allowed modulation in the 5 GHz range and used orthogonal frequency-division multiplexing (OFDM) to let one device use multiple carriers at the same time. 802.11b remained in the 2.4 GHz range and used direct-sequence spread spectrum (DSSS) modulation (like the original standard) but increased the maximum theoretical transfer rate to 11 Mbps. The 802.11g standard, which operates in the same 2.4 GHz range, was released in 2003 and provided a significant upgrade from ‘b’. 802.11g uses OFDM to increase the maximum theoretical data rate to 54 Mbps but is also backwards compatible with 802.11b hardware [13]. The next PHY upgrade, seeking to improve transfer rates of the ‘a’ and ‘g’ amendments, was the

802.11n amendment which was released in 2009. 802.11n was the first 802.11 specification to standardize multiple-input multiple-output (MIMO). This standard supports up to four different data streams and also allows the use of 40 MHz channels, which is a significant change from the previous designation of 20 MHz channels. 802.11n was specified to operate in either the 2.4 GHz or 5 GHz range and soon became popular in the marketplace. It allows for a maximum theoretical data rate of 72.2 Mbps per data stream for 20 MHz channels and 150 Mbps per stream for 40 MHz channels [13], though practical speeds in a commercially deployed WLAN environment are much lower. Finally, the most recent PHY upgrade to become commercially available is 802.11ac, which is designated to operate in the 5 GHz range and allows for 20, 40, 80, or 160 MHz channels. The widest channel allows a maximum theoretical speed of 867 Mbps per data stream [14], but only rarely would a similar rate be achieved in practice.

IEEE has also updated its MAC specifications over the years, though perhaps with not as much fanfare to the consumers as the PHY specifications. The Wired Equivalent Privacy (WEP) encryption scheme was shown to be vulnerable to malicious attacks, so 802.11i introduced a much more secure mechanism called WPA2 (the name derived from a market-developed solution called Wi-Fi Protected Access, or WPA) [13]. Other notable MAC amendments were 802.11e which addressed different quality of service (QoS) classifications (among other things), 802.11z which solidified specifications for direct link between devices without an access point (AP), and 802.11u which specifies interworking with external networks.

The fact that WLAN APs have smaller communication ranges than most cellular base stations is both a benefit and a drawback. On the one hand, a relatively large number (compared to cellular deployments) of Wi-Fi APs can be placed in densely populated areas to service a larger data traffic demand due to increased frequency reuse. On the other hand, mobile WLAN access is more difficult as the AP coverage is smaller and disconnections are more frequent. In the context of immediate data usage, though, currently available Wi-Fi devices provide higher maximum theoretical data rates than cellular communications and they generally have better throughput performance in a practical metropolitan environment [15].

2.1.2 Femtocells

A femtocell is a low-power cellular base station intended to cover a very small area (such as inside a house). Its purpose is similar to that of a Wi-Fi AP in that it allows for localized network access and can be used to increase capacity and improve coverage in a densely populated area. It has its own backhaul connection (not necessarily coupled to the MNO's backhaul), but it uses cellular protocols and the same licensed frequency bands used by MNOs, so it appears to a mobile device as a traditional cellular base station. Similar to WLANs, femtocells are considered a viable outlet for offloading data from cellular networks.

A unique benefit of femtocells is that they are necessarily managed by the cellular carrier and therefore can usually provide more seamless network access. This is in contrast to

WLAN access, which apart from carrier deployment can cause an interruption in services that require a smooth handoff. On the other hand, this necessary coupling of femtocells to the MNO's network can sometimes cause signaling and traffic delay, as the femtocell's backhaul connection may not support the same QoS requirements as the MNO's backhaul [16]. Another drawback of femtocells is that they sometimes cause interference with the macro/micro cellular infrastructure, particularly for cell-edge users when the femtocell is under closed access (only registered devices may connect).

There is a good deal of research specifically related to femtocell offloading. Though there are many parallels between WLANs and femtocells, and sometimes offloading models for one can be roughly applied to the other, this thesis focuses mainly on the challenges and solutions specifically related to data offloading using WLAN access.

2.2 History and Current State of Wi-Fi Offloading

Throughout the world, mobile providers who wanted to offload data from their cellular networks using public WLAN hotspots were initially limited to deploying their own Wi-Fi networks. Some companies, such as AT&T and China Mobile, made substantial progress. China Mobile in particular has made a robust effort at deploying Wi-Fi hotspots in dense areas and by 2013 reported that half of all their mobile traffic is delivered by these Wi-Fi APs [17].

For a more detailed look at how an MNO has been able to shift traffic from cellular to Wi-Fi, consider AT&T's efforts in the United States. AT&T for some time had a reputation of poor mobile network performance, due in large part to their relationship with Apple [18] (AT&T was the sole carrier of the iPhone for three years). The sharp increase of smartphone customers combined with unlimited data plans (for a time) spelled disaster for their mobile network.

In addition to making massive expenditures to upgrade their 3G network, AT&T was concurrently deploying their own Wi-Fi hotspots nationwide to alleviate traffic demand, particularly at chain restaurants and coffee shops. Even as far back as 2008 they had deployed 17,000 hotspots [19], though the offloading effect was not so great as to have a significant impact on their network congestion. Even with arguably the widest Wi-Fi network in the country AT&T still experienced significant cellular network congestion, indicating that their Wi-Fi infrastructure was not yet effective as a significant offloading tool. AT&T continued expanding their Wi-Fi network coverage and functionality. In 2012 they saw a 200% increase in mobile traffic on their Wi-Fi network from 2011 [20]. Though there's little information offered to indicate whether this spike is due in larger part to an increase in overall demand or a more far-reaching and effective Wi-Fi network, this information indicates that AT&T and its subscribers are both looking to Wi-Fi as a useful and even preferable alternative to standard cellular data.

With Wi-Fi becoming an increasingly valuable offloading choice, the landscape for Wi-Fi deployment is changing. Boingo Wireless, a commercial Wi-Fi provider, initially provided

Wi-Fi hotspots in airports and hotels for individual subscribers, particularly business travelers. In the past few years, though, they have been solidifying roaming agreements with a number of large wireless providers, many of whom having their own Wi-Fi deployments. Even before Hotspot 2.0 (explained in the next section) was commercially available, Boingo was known to have roaming agreements with AT&T [21], SK Telecom [22], NTT Docomo [23], China Telecom [24], and others. Now that Hotspot 2.0 *is* beginning to be deployed, Boingo's prospects are likely to improve even further. In fact, they've already signed an offloading deal with Verizon [25] in the hopes that Hotspot 2.0 will soon be beneficial to both parties. This increased market penetration has allowed Boingo to expand their business focus to cover not just airports and hotels, but also places like coffee shops and even the New York City subway [26].

Though Boingo Wireless has had success in sealing roaming agreements, it is only one of several non-mobile network providers who have entered into the Wi-Fi hotspot business. Among others, Comcast, Time Warner, and KT (Korea) are wireline internet providers who have deployed their own Hotspot 2.0 enabled Wi-Fi hotspot networks. All three of these are among the Wireless Broadband Alliance's (WBA) Next Generation Hotspot (NGH) program participants, who have been involved in the WBA's testing of the NGH architecture (explained in the next section). NGH technology was first demonstrated to be commercially ready by the WBA at a meeting in China in October, 2013 [27], and since then several wireless and wireline providers have begun to deploy Hotspot 2.0 enabled hotspots. This appears to be a pivotal time for the emergence of large-scale Wi-Fi roaming.

2.3 Hotspot 2.0

The IEEE 802.11u (from here on referred to as “11u”) amendment, titled “Interworking With External Networks”, is important for the future of Wi-Fi offloading. It provides MAC layer specifications that facilitate seamless interworking and looks to be beneficial for mobile users. Implementation of the new standard only requires a firmware upgrade, which avoids the infrastructure cost of purchasing new hardware for service providers who want to deploy these networks.

Under the 11u standard, the AP can now include information elements in its beacon frame and/or probe response frame that indicate whether they support interworking and a roaming consortium [28]. This way a mobile device can learn which Wi-Fi networks support a roaming agreement with its mobile carrier before association, and therefore can prioritize the order of APs with which it tries to authenticate. Once the client has determined this prioritized list, the subsequent association requests/responses involve exchanging interworking information for the purpose of authentication.

Hotspot 2.0 is an initiative by the Wi-Fi Alliance with the purpose of creating industry standards that would improve the Wi-Fi hotspot experience by making it “as secure and easy to use as cellular” [29]. The following is an example of the communication between a cellular customer’s mobile device and several WLAN access points as dictated by the Hotspot 2.0 framework. The user equipment (UE) receives beacon information from the Wi-Fi APs and can probe request these APs using Generic Advertising Service (GAS), allowing

it to determine from the beginning which APs support interworking. As many as three Organizational Identifiers (OIs are indicative of particular members in the SSO's roaming consortium) can be broadcasted by the AP in the beacon frames and probe responses [28]. If the UE needs to, it can use the Access Network Query Protocol (ANQP) to communicate further with a particular AP in order to find out the full list of supported external networks. The UE compares this information from the several APs and determines the most appropriate option.

After this process is complete, the UE then requests association with the preferred access point and attempts to authenticate via a pre-provisioned mechanism. In the case of a cellular user on a GSM network, the mechanism will likely include a submission of SIM (subscriber identity module) credentials. The AP, which is itself pre-provisioned to be able to contact servers of providers in its roaming consortium, forwards the SIM credentials to the proper authentication server. When the AP receives a confirmation of authorization it then allows the UE to connect to the Wi-Fi network. This entire process can be performed in seconds or less with no participation from the user.

Building on the 11u standard and the Hotspot 2.0 framework, the WBA is leading the NGH project, which aims to establish interoperability among service providers. This initiative can be seen as the actual implementation of Hotspot 2.0 on real-world networks. NGH is developing methods of interconnecting networks for easy roaming authentication in order to achieve an effective and secure Wi-Fi experience. The three primary components of this are the 11u standard, WPA2-Enterprise WLAN security, and EAP-based (Extensible

Authentication Protocol) authentication [29]. Different devices may use different forms of EAP authentication. For instance, a mobile phone on a GSM network would use EAP-SIM or EAP-AKA (Authentication and Key Agreement) for a legacy GSM network or a UMTS (Universal Mobile Telecommunications System) network, respectively. UEs on CDMA networks may use a pre-approved username/password combination, which would require EAP-TTLS (Tunneled Transport Layer Security). Finally, those devices wanting to use a X.509 certificate would use EAP-TLS (Transport Layer Security).

The Hotspot 2.0 framework is now commercially deployed but still in its infancy. Though a relatively small number of mobile providers worldwide have implemented this technology and established the needed roaming relationships, its use is expected to grow. The WBA has predicted that by 2018, more than 80% of public Wi-Fi networks will have fully automated roaming functionality [17]. It is this type of worldwide Wi-Fi roaming that we're expecting can benefit from an offloading model such as the one presented in this thesis.

2.4 Device Handoff

In the context of offloading mobile network traffic to WLAN networks, there are three philosophies on how close the service remains to the MNO's backhaul: tight coupling, loose coupling, and no coupling. Tight coupling implies that data is being routed to the user via the MNO's core network, whereas loose coupling indicates that the data transfer is apart from the MNO's backhaul but there is an interworking element where authentication is shared.

Finally, no coupling merely implies that a user is manually connecting to an independent Wi-Fi AP and none of his traffic is routed through the MNO's core network.

Offloading with no coupling can be achieved any time a user has access to an independent Wi-Fi AP. This has been available ever since mobile phones began to have Wi-Fi radios. Moving forward, the industry is pursuing solutions that require coupling but are also more easily accessible to the user. The use of Hotspot 2.0 enabled networks is an example of loose coupling, where data is routed through the IP network but authentication is performed strictly through the MNO's network. WLAN sessions with no coupling or loose coupling are almost always data-centric and do not require significant additional attention to handoff mechanisms. While not ideal, packet-switched services can often afford to suffer minor latency delays during Wi-Fi to Wi-Fi horizontal handoff or vertical (cross-medium) handoff between 3G/4G and Wi-Fi without significant inconvenience to the user. Call-centered applications, on the other hand, usually require tight coupling and effective handoff techniques, both in Wi-Fi to Wi-Fi handoff as well as in vertical handoff.

In order to ensure effective handoff for phone calls or other time-critical sessions using Wi-Fi we need to examine the technology in place for each scenario. If two Wi-Fi APs are in the same extended service set (ESS) but require the 802.1X authentication process then the handoff between the two could be relatively lengthy and cause a dropped session. This handoff is made simple under the "fast transition" provision of the 802.11r standard [30]. The "11r" amendment allows the UE negotiate a security key and request wireless resources at the same time, thus shortening the process. The only safeguard a network administrator

must provide is that the APs are positioned close enough to each other that the UE travelling between the coverage area of the two of them is never too distant that it can't communicate with an AP at all.

In the case that the UE is travelling through overlapping coverage areas of two Wi-Fi APs that are not part of the same ESS, the authentication process with a new AP may take long enough that a data session suffers or is dropped completely. This potentially lengthy transition could be addressed by a mobile IP solution. This involves a device having both a home IP address and a "care-of address" (CoA) which is associated with the network to which it's currently connected. Information is sent to the network agent associated with the home address, which in turn tunnels the packets to the CoA.

In the case of a vertical handover, the 802.21 standard for Media-Independent Handover (MIH) is more applicable. This standard involves specifications on how the link layer needs to communicate with higher level layers. For instance, if a session is using a Wi-Fi connection and the signal is weakening, the link layer relays this message to the Media Independent Handover Functions (MIHF) operating on a higher protocol layer on the UE. This message is in turn relayed to the MIHF on the AP, which then gives a command back to the UE to measure signal strength on all APs (Wi-Fi, 3G, 4G, etc). The UE assesses the best connection and petitions the AP of that connection for performance information. This new AP will ensure that it can allocate resources and then instruct the UE to commit to the MIH in order to begin receiving data on the new connection.

Unfortunately the MIH process requires a great deal of communication and synchronization, so it is never guaranteed to be perfectly effective. If an MNO is establishing an infrastructure for WLAN offloading it would probably be in its best interest to make WLAN connection a preference only for non-critical data-centered applications and retain the cellular network as a priority for voice calls and other time-critical sessions. Perhaps they can provision their networks such that calls are allowed to be conducted over WLAN only in the event that no cellular connection is available.

2.5 Delay-Tolerant Data and Opportunistic Offloading

Most of the WLAN offloading that has been achieved in practice has focused on real-time data. Mobile subscribers largely use data at their own pace, and though they may choose to wait until connecting to a Wi-Fi hotspot before engaging a data-heavy application, the fact that this opportunistic decision is made entirely by the user reveals an inefficiency. In response to this there have been substantial research efforts regarding the delay-tolerant nature of certain data applications. For example, a video upload or a data backup could require a relatively large amount of mobile data resources compared to other data sessions, but these two events are also generally not time critical. If a video upload were to be afforded some maximum delay time (perhaps 10 minutes) after the action is cued by the user, there is a chance that a Wi-Fi hotspot would be available before the deadline expires, and therefore the data session could possibly be conducted over Wi-Fi instead of the cellular network. The

most common topics regarding delay-tolerant offloading are the systems required to enact opportunistic offloading that is invisible to the end-user, including extending the principle of opportunistic offloading to peer-to-peer connections using an ad hoc network, and the effectiveness of opportunistic offloading, including the relationship between the offloading deadline and the percentage of delay-tolerant data that can be offloaded (in a specific wireless environment).

In order to take advantage of opportunistic offloading at all, there must be a proper framework in place to allow such events. Usually delay-tolerant networks (DTNs) require some form of “store-carry-forward” capability [31] for each device in the network out of necessity. In the present case though, we intentionally want to delay the transmission (though the cellular data network is immediately available) until a WLAN connection is available.

The authors of [12] propose a framework called MADNet that is based on the cooperation of MNOs, SSOs, and separate user devices. When a user requests content via a cellular connection, the BS would be aware of nearby Wi-Fi APs and would make a decision about whether or not to delay information delivery. Alternatively, content could be sent over the cellular connection to a particular few users, who would then disseminate the information to nearby UEs who need the same content. The authors of [32] explore the use of ad hoc opportunistic communication more deeply by establishing a more complex scenario with different types of data streams (similar to a real-world network) and optimizing the amount of data offloaded before the assigned deadline.

More relevant to the topic of WLAN offloading, however, is the work presented in [33], where the authors specify that a device connected to an infrastructure network would first attempt to send information over the DTN. They used their own routing scheme (previously outlined in [34]) which allows the sending device to be aware of the receiving device when the links are in place for the two to successfully communicate. The original sending device would only use its infrastructure-based network if it considered the probability of the message reaching its destination through the DTN low. Though this is not a perfect parallel to intermittent Wi-Fi access, the idea can be applied: A UE that needs to send or retrieve content would wait until it establishes a connection with a Wi-Fi AP. If the message is large and the WLAN connection is tenuous, the UE may decide that it is unlikely that the content will be delivered and resort to using cellular network resources.

One of the most commonly cited works regarding the effectiveness of delay-tolerant offloading addresses the case of data use in moving vehicles through a city with limited Wi-Fi coverage [11]. The paper presents a novel system called Wiffler that takes advantage of delay-tolerance and selects Wi-Fi as a preference over 3G, but also has a fast-switching capability to return from Wi-Fi to 3G in the event that the UE moves out of range of the WLAN. Despite having access to a Wi-Fi AP only 11% of the time, Wiffler was able to reduce 3G usage by 30% when performing 5 MB transfers with a maximum delay tolerance of 60 seconds.

Another prominently cited paper paints a different picture of the effectiveness of delay-tolerant offloading. The authors of [35] performed a study that tracked the movement of 97

iPhone users in South Korea to obtain information on their daily network access, including the proportion of time during which the users had Wi-Fi access as well as the data speeds of the various connections (3G vs varied WLAN APs). The authors then performed a simulation of mobile data sessions using the Wi-Fi AP information they had previously collected. They concluded that “on-the-spot” offloading (immediate data usage while Wi-Fi happens to be available) can already account for 65% of the data demand of these users and that adding a 100 second delay tolerance only accounts for an additional 2%-3% of the total data being offloaded. It was only when the delay tolerance was an hour that the additional offloading gains reached 20%. This result may seem pessimistic but there are certainly applications that can tolerate an hour’s delay (such as large file transfers), which could still render a delay tolerant system very useful.

In addition to the necessary architectures and the effectiveness of such schemes, some people are considering the energy efficiency aspect that a 3G radio consumes more power than a Wi-Fi radio. The authors of [36] present a framework such that when a one hour delay is tolerated, in addition to more than 20% of traffic being offloaded to WLANs, between 20% and 35% less energy is used by the mobile application. These numbers improve even further with greater delay tolerances.

Delay-tolerant offloading continues to be an active research area. While some findings are not as encouraging as others, the body of research performed in this area reveals potential for application in mobile devices. Despite these advances, there are few applications or platforms in common use that take advantage of the delay-tolerant nature of certain types of data.

Usually the delay tolerance would have to be incorporated into a particular application on a phone, and it's likely that few developers currently consider this a feature in demand by mobile subscribers. Regardless, as mobile data consumption continues to increase and mobile users choose smaller data plans, the notion of delay-tolerant offloading has the potential to become more appealing.

2.6 Economic Models for Offloading

In order to make an offloading system effective in an industry deployment there must be a pricing model for the service offered, such as the one explored in this thesis. Pricing models for network access have been the subject of a great deal of research. Some efforts focus on time-dependent pricing to decrease demand at peak times rather than servicing the entire demand using other outlets [37]. In this thesis, though, we consider pricing models in which a provider leverages a separate network's resources (or an access network other than the provider's standard service) in order to service the increasing mobile data demand. One early pricing model of this type employed a cellular/Wi-Fi heterogeneous network with volume-based pricing over GPRS and flat pricing over Wi-Fi and analyzed a user demand function that depends on these two charges [38]. As commercial WLAN access has become more common and more able to service offloading needs, however, many different complex offloading models have emerged.

Building on the topic discussed in the previous section, we first discuss the pricing models that focus on offering incentives to mobile users to leverage the delay tolerance of their data. The authors of [39] present a framework based on a reverse auction in which mobile users place bids to the MNO with their delay tolerance and desired discount. The MNO then prioritizes the users' bids and delays certain users' traffic delivery within their stated delay tolerance in a way that minimizes the total discounts that they must deliver. Another aspect of delay-tolerant pricing focuses on quantifying the actual economic benefits to both the MNO and mobile subscribers when such an offloading framework is in place. The authors of [40] present a game theoretic pricing framework where the MNO sets prices and the user accepts the price. They find that depending on the pricing scheme (e.g. flat, volume-based, etc.) the MNO can increase its revenue and a user can increase its surplus, defined as the difference between the price the user is willing to pay and the price he does pay.

We also want to consider more general pricing models that don't necessarily require delay tolerance. The authors of [41] present a model of femtocell offloading similar to the one presented in this thesis with established roles of large-scale and small-scale operators. This work considers the relationship between the price charged for accepting a foreign user and the resources set aside for accepting foreign users. Similarly, [42] explores the challenges of overlapping WiMAX and Wi-Fi coverage, where the two radio access technologies are part of separate networks run by separate operators. The author uses a genetic algorithm for the two operators to determine each other's bandwidth demand in order to set a price. Both of these models are related to the one presented in this thesis, in which one operator

considers the user demand and cost structure of the other operator's network in order to set an offloading price that will maximize the utility of both parties.

The key difference between the aforementioned offload pricing models and the one presented in this thesis is the party making the decision. In the model presented here, each MNO decides how many of its subscribers it wants to offload, rather than allowing its subscribers to choose for themselves. We examine the situation in which a user does not have free access to existing Wi-Fi networks and must rely on their MNO to form roaming agreements with private SSOs. When the MNO must pay for the offloading service, it has a vested interest in controlling the number of subscribers it allows to be offloaded. Therefore the pricing mechanics and offload decision is entirely based on the interaction between the MNO and SSO, and is transparent to the end-user.

The authors of [43] present a model more similar to the one in this thesis, in that the MNO makes its own offloading decisions, though the transaction process differs as it calls for a price to be offered by the MNO to the APs' owners (the APs may not be aggregated into a single network). The reverse of this, where the AP owners offer prices to the MNOs, is presented in [44].

We also consider game-theoretic offloading models that demonstrate competition exclusively between operators (and not the end-user). One such model, presented in [45], focuses on power allocation to maximize data rate. Though it does not address monetary pricing, it does share with the model presented in this thesis the principle of making an offloading decision based on optimizing a utility function that is affected by the actions of other service

providers. Another game-theoretic model in which the offloading decision is transparent to the mobile subscriber is presented in [46]. This model shares with ours the concept of MNOs engaging in incentive-based relationships directly with Wi-Fi AP operators, but once again differs in that the MNOs are the parties initiating the price.

In addition to models strictly related to offloading data from one network medium to another, it is also helpful to examine pricing models for Mobile Virtual Network Operators (MVNOs). An MVNO is a mobile operator that has no network of its own but instead services its subscribers by buying limited access to another MNO's network. Often times these operators will offer prepaid phone options and less expensive plans. There is a parallel between the operation of an MVNO and an MNO in the offloading model presented in this thesis. Though an MNO in this thesis has its own network on which it attempts to service most of its subscriber base, it still must purchase access to another network's resources in order to service a substantial portion of users. Just as this traditional MNO offloads some of its subscribers onto a different network, an MVNO can be said to "offload" its entire subscriber base onto a different network.

One MVNO model addresses the dilemma of the MNO who is selling network access regarding how much bandwidth to sell [47]. If the MNO sells too much bandwidth it will lose money as a result of suffering poor QoS for its customers, but if it sells too little bandwidth it will end up with unused resources, which can also represent lost income. From this perspective, the authors use game theory to maximize the MNO's revenue. If we were to use this scenario as an analogy to the offloading relationship presented in this thesis, the

SSO would be the party concerned with how much bandwidth it is allowing the MNO to access. In fact, the model presented here includes in the SSO's utility function a term that serves as a hedge against the QoS burden levied by the MNO's offloading.

The pricing models discussed above span a wide variety of network sharing/offloading scenarios. The model presented in this paper draws on key themes addressed in these papers but ultimately provides a novel approach to the economic relationship between two operators.

2.7 Viability of WLAN Offloading as a Long-term Solution

Though the vast amount of research and industry advancement on WLAN offloading presents an optimistic outlook, there remain doubts about whether it is a viable long-term solution to the increasing mobile data demand. For one, WLANs have historically been considered far less secure than 3G and 4G cellular networks and have been labeled “untrusted” by mobile providers [10]. Security concerns have largely been addressed by advancements in relevant standards such as IEEE 802.1X authentication and the encryption provided by IEEE 802.11i. Furthermore, the advent of shared authentication via Hotspot 2.0 makes it more likely that mobile operators will consider it a “trusted” technology [10].

Another key indicator of WLAN's viability as a long-term solution is how much traffic it is actually capable of offloading. Previously discussed results for delay-tolerant offloading

indicate that Wi-Fi offloading can significantly reduce the load on cellular networks. When the delay tolerance is low (e.g. 1 minute), one paper claims that the vast majority of offloaded data is not even delayed at all, but rather is accomplished by simple “on-the-spot” offloading [35] which is already widely deployed in practice. Additionally, there is evidence (outlined previously) that efforts by providers such as AT&T and China Mobile to deploy carrier Wi-Fi have resulted in substantial proportions of their mobile traffic being carried over WLANs.

Finally, the perceived tendency of industry to embrace Hotspot 2.0 seems to indicate that WLAN offloading is a highly desired outlet for mobile data. The USA’s largest cellular carrier, Verizon Wireless, had never made a significant effort at deploying carrier Wi-Fi or otherwise offloading traffic onto WLANs until they signed a roaming agreement with Boingo Wireless in 2014 that takes advantage of Hotspot 2.0 APs. Furthermore, wireline Internet providers such as Comcast, Time Warner, and KT have indicated their commitment to deploying their own Hotspot 2.0 APs, likely with the expectation of engaging in lucrative roaming deals with cellular carriers. These industry moves could be indicative of a network access paradigm shift to a system that relies on WLAN access.

2.8 Conclusion

The advancement of Wi-Fi as a trusted technology along with the vast amount of research conducted on Wi-Fi offloading point to a future with increased mobile WLAN access on a

daily basis. With this in mind, we now explore our own offloading model with which we seek to learn about the nature of the MNO/SSO offloading relationship.

Chapter 3

Offloading Model and Primary Exploration

In this chapter we explore the offloading relationship between MNOs and an SSO. We begin by specifying the model in use and then expanding with mathematical analysis and results. In this model, the SSO sets an offloading price per person and the MNO responds by offloading a certain number of users, depending on its own expenses. The SSO anticipates the way the MNO will behave and accordingly sets a price that maximizes its own utility.

In addition to having its own direct implications, we find that this model helps us better

Most of the content in this chapter is also included in a paper that is under review for publication in IEEE Transactions on Mobile Computing as an extension of [1], ©2014 IEEE.

understand the relationship between an SSO and several MNOs with different cost function parameters that all want to offload subscribers.

3.1 Model and Problem Overview

3.1.1 Notation and Fundamental Relationships

Let M be the total number of MNOs that have an offloading agreement with the SSO. Each MNO i has a cost function denoted by $\psi_i(n)$ which is a function of the number of currently active users on the cellular network (this function does not include offloading costs). Each MNO also has several important parameters to consider:

- $n_{mno,i}$ is the total number of active subscribers for which MNO i must provide service, whether on its own network or offloaded to the SSO's network.
- χ is the price set by the SSO to offload a single user.
- n_i is the number of users that MNO i will choose to service natively (on its own network) once the SSO sets an offloading price.
- $n_{off,i}$ is the number of subscribers that the MNO chooses to offload to the SSO's WLAN network as a result of the SSO setting an offloading price. Note that $n_{off,i} = n_{mno,i} - n_i$.

The MNO's utility function is denoted by,

$$U_{mno,i} = - [\psi_i(n_i) + n_{off,i}\chi]. \quad (3.1)$$

This utility function includes two primary costs: Servicing users on the MNO's native network (ψ_i term) and offloading users to the SSO's network. Inherent in the MNO's cost function ψ_i is a consideration of the QoS burden that the MNO suffers on its native network as the number of active subscribers increases.

The SSO's utility function is denoted by,

$$U_{ss0} = n_{off}\chi + d_1 n_{ss0} \log \left(\frac{E}{n_{ss0} + n_{off}} \right), \quad (3.2)$$

where:

- n_{off} is the number of users offloaded from all MNOs, and therefore currently being serviced by the SSO's network,
- n_{ss0} is the number of native (non-offloaded) users on the SSO's network,
- E is the SSO's total bandwidth, and
- d_1 is a scaling factor to establish an appropriate relationship that defines the tradeoff between the revenue due to the offloaded users and the strain that those users place on the SSO's network.

The first term in the SSO's utility function is merely the revenue earned from the MNOs, which is directly proportional to the total number of offloaded users. The second term is intended as a hedge against the QoS burden that the MNOs' offloaded users place on the SSO's network. The positive \log term with the average bandwidth per user as its argument

(or equivalently, the negative *log* term whose argument is directly proportional to the total number of users on the SSO's network) captures the diminished marginal SSO utility as the number of offloaded users increases.

3.1.2 MNO Cost Function

We consider the cost function ψ_i of each MNO to be indicative of the monetary expenses of infrastructure investment as well as the less tangible costs of the need for network expansion and the potential loss of customers in case QoS deteriorates.

Using this approach we consider that each MNO has a piecewise cost function with two regions, as depicted in Figure 3.1. In the first region, from 0 to a certain threshold number of users $n_{T,i}$, the cost function is constant. This represents the fixed infrastructure costs of the initial network deployment which can service up to $n_{T,i}$ users before being capacity-constrained. In other words, there is zero marginal cost of serving a subscriber, up to $n_{T,i}$ users.

The second region, which begins immediately following the threshold $n_{T,i}$, is characterized by a strictly convex increase. In this region the MNO becomes increasingly capacity-constrained and the performance of its network begins to suffer as the number of active subscribers increases. The cost function increase represents higher operating costs as well as the need to make further network investments to better service the current users. Furthermore, we consider that the risk of losing customers becomes more significant. In many

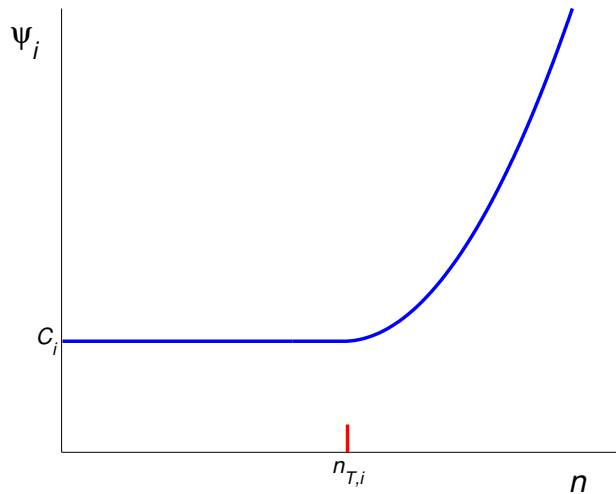


Figure 3.1: MNO cost function with a strictly convex increase after surpassing a threshold number of users.

wireless telecom markets the primary factors for customer loyalty are the satisfaction and trust in the MNO’s network performance [48] [49], and customer loyalty is extremely important for companies in service-oriented industries to maintain a competitive advantage [50].

3.1.3 Behavior of an Individual MNO

When an MNO is operating in the strictly convex region of its cost function (past its threshold $n_{T,i}$), we consider the marginal cost for the MNO to service the n^{th} additional user on its native network to be approximately equal to $\psi'_i(n)$, which is the derivative of its cost function at n users. In response to the SSO’s per-user offloading price, the MNO will only offload a user if the cost to do so is less than the marginal cost of servicing that user on its native network. In the “indifferent” case that the price is the same, we assume that the MNO

chooses to retain this user on its own network. Formally, the MNO will offload user n if $\chi < \psi'_i(n)$.

Since the MNO's cost function has a strictly convex increase, this also implies that every single user past n up to $n_{mno,i}$ will also be offloaded. This is in keeping with the constraint that $n_{off,i} = n_{mno,i} - n_i$ (the number of offloaded users is equal to the difference between the total number of active MNO subscribers and the number of users serviced by the MNO's native network).

In practice the SSO's price χ will fall within an interval $(\psi'_i(n_i), \psi'_i(n_i + 1)]$ for each MNO i . However, for simplicity of mathematical manipulation, we will consider the offloading price to be approximately equal to the derivative of each MNO's cost function at its respective value of n_i . Formally, we will use the following relationship in our analysis for each MNO i :

$$\chi = \psi'_i(n_i) \tag{3.3}$$

We have, then, that n_i equals the inverse of the derivative of ψ_i at χ , which we denote as $h_i(\chi)$:

$$n_i = \psi'^{-1}_i(\chi) = h_i(\chi)$$

Finally, we can express the number of users offloaded by an individual MNO as:

$$n_{off,i} = n_{mno,i} - h_i(\chi). \tag{3.4}$$

3.1.4 Problem Outline

We wish to determine the optimal offloading price the SSO must charge in order to maximize its own utility. For the majority of this thesis we assume that the SSO has complete information of the MNOs' cost structure and utility. (We address the absence of complete information in our discussion of the results.) Using the argument outlined in the prior section along with specific cost function parameters, the SSO can determine how many users each individual MNO will offload at a given price. This exploration considers a single moment in time where the number of active MNO subscribers is constant. We leave for future work a time-evolving analysis in which each MNO has a varying active user base.

In the scenario where many MNOs have an offloading agreement with the SSO we will find that, based on the varying cost functions of the MNOs that are offloading, different MNOs will offload at different price ranges. If the price were to continually fall from an initially high value we would observe that each individual MNO would begin to offload its respective users at a particular price which is likely to be unique among all the MNOs (we explore this effect in subsequent sections). For this reason, we split up the SSO's decision space for its offloading price into $M + 1$ regions, as seen in Figure 3.2. The index of each region corresponds to the number of MNOs that will offload at least one user for all prices in that region. In other words, region k encompasses all values of the price χ that will entice exactly k MNOs to offload at least one user. Furthermore, we index the MNOs according to the region in which they begin to offload. Region 1 sees offloaded users from MNO 1, region 2 from MNOs 1 and 2, region 3 from MNOs 1, 2, and 3, and so forth.

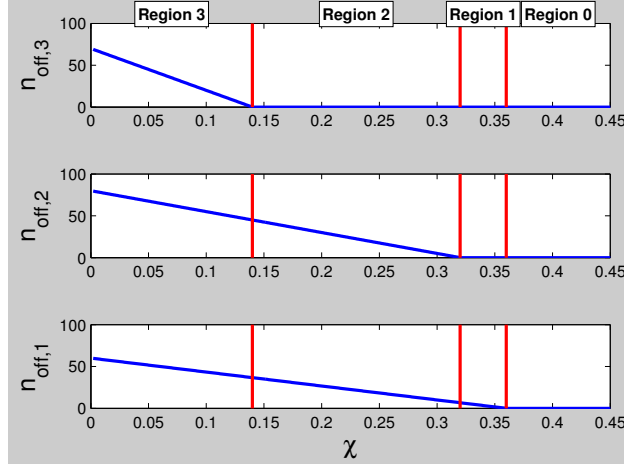


Figure 3.2: The SSO's decision space for setting the offloading price χ is split up into $M + 1$ regions.

We must identify the SSO's locally optimal price for each decision region k and then determine which of these prices maximizes its utility over the entire decision space. To begin this process, we first express the number of users offloaded from all MNOs for a particular region k as a function of the SSO's offloading price. This value, denoted by n_{off}^k , can be found by summing the RHS of Equation 3.4 for each MNO that offloads users in region k :

$$n_{off}^k(\chi) = \sum_{i=1}^k (n_{mno,i} - h_i(\chi)) \quad (3.5)$$

We can substitute this relationship into the SSO's utility function and differentiate to find the value of χ that maximizes the SSO's utility within region k .

A locally optimal price that maximizes the SSO's utility within decision region k can be represented by $\chi_{k,U}^*$, and the globally optimal price is the one that maximizes the SSO's

utility over all values of values of χ :

$$\chi_U^* = \arg \max_{k \in \{0,1,\dots,M\}} U_{ss0}(\chi_{k,U}^*). \quad (3.6)$$

3.2 Piecewise-Quadratic Cost Function

For concreteness, we consider the case in which each MNO has a piecewise-quadratic cost function:

$$\psi_i(n) = \begin{cases} C_i & n \leq n_{T,i} \\ \alpha_i(n - n_{T,i})^2 + C_i & n > n_{T,i} \end{cases}. \quad (3.7)$$

3.2.1 Number of Users Offloaded

Equation 3.4 tells us the number of users that a particular MNO i will offload in terms of the MNO's cost function and the offloading price that the SSO charges. Now that we've specified the cost function, we can follow the procedure set forth in Section 3.1.3 and derive an explicit function for $n_{off,i}$.

$$\begin{aligned} \chi &= \psi'_i(n_i) \\ &= 2\alpha_i(n_i - n_{T,i}) \\ n_i &= \psi_i'^{-1}(\chi) \\ &= \frac{\chi}{2\alpha_i} + n_{T,i} \end{aligned}$$

Therefore, for a particular MNO i ,

$$n_{off,i} = b_i - \frac{\chi}{2\alpha_i}, \quad (3.8)$$

where $b_i = n_{mno,i} - n_{T,i}$. Using this result along with Equation 3.5, we can determine the total number of users offloaded onto the SSO's WLAN by all MNOs that are active in a given decision region k :

$$\begin{aligned} n_{off}^k &= \sum_{i=1}^k \left(b_i - \chi \frac{1}{2\alpha_i} \right) \\ &= \underbrace{\sum_{i=1}^k b_i}_{B_k} - \chi \underbrace{\sum_{i=1}^k \frac{1}{2\alpha_i}}_{D_k} \\ &= B_k - \chi D_k. \end{aligned}$$

3.2.2 Simplified SSO Utility: Revenue Only

In order to gain intuition, we first consider the scenario where the SSO is only trying to maximize its revenue, represented by G . The SSO's revenue is simply the product of its offloading price and the number of users offloaded:

$$G = n_{off}\chi. \quad (3.9)$$

After the SSO sets its offloading price we can express its utility curve for a particular decision region k (where k MNOs choose to offload users) by the following:

$$\begin{aligned} G_k &= (B_k - \chi D_k)\chi \\ &= -D_k\chi^2 + B_k\chi. \end{aligned}$$

Knowing this relationship, the SSO finds the price that maximizes its revenue by simply taking the derivative of G_k and setting it equal to zero. This optimal price is:

$$\chi_{G_k}^* = \frac{B_k}{2D_k}. \quad (3.10)$$

It is important to understand that this price maximizes the curve seen in region k of the SSO's full piecewise revenue relationship, but does not necessarily represent the price that maximizes the revenue *within that region*. As an example, consider Figure 3.3. The solid piecewise curve is the revenue that the SSO will earn for a given price, taking into account the different numbers of MNOs that are offloading in each region. The dotted curve is the continuation of the curve seen in region 1 (where only one MNO chooses to offload). That is, the dotted curve represents what the SSO's price-revenue relationship *would* be if the MNO choosing to offload in region 1 were the only MNO that had an offloading relationship with the SSO. In summary, this is the generalized notation:

- $\chi_{G_k}^*$ is the offloading price that maximizes the continuous extension of the curve seen in region k , as though the SSO has no other decision regions.
- $\chi_{k,G}^*$ is the price that maximizes the SSO's revenue within region k .
- χ_G^* is the price that maximizes the SSO's revenue over its entire decision space.

We see here that the locally optimal price $\chi_{k,G}^*$ will either be $\chi_{G_k}^*$ or one of the two values of χ that form the boundaries of region k .

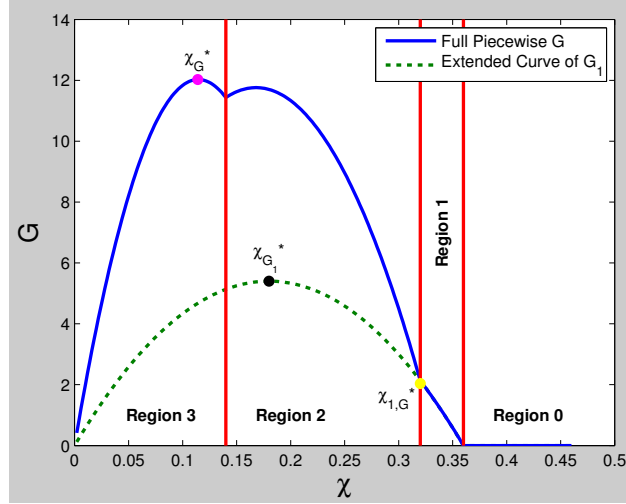


Figure 3.3: Difference between $\chi_{G_k}^*$ and $\chi_{k,G}^*$.

Formally, the price that maximizes the SSO's revenue within region k can be expressed

by:

$$\chi_{k,G}^* = \theta \cdot \chi_{G_k}^* + (1 - \theta) \left[\arg \max_{\chi \in \{\chi_{k,min}, \chi_{k,max}\}} G(\chi) \right], \quad (3.11)$$

where

- $\chi_{k,min}$ is the minimum value of χ that falls in region k , which is equal to the value of χ that is the boundary between region k and its adjacent region with lower values of χ .
- $\chi_{k,max}$ is the maximum value of χ that falls in region k , which is equal to the value of the χ that is the boundary between region k and its adjacent region with higher values of χ .

- θ is an indicator variable with the following definition:

$$\theta = \begin{cases} 1 & \chi_{G_1}^* \in [\chi_{k,min}, \chi_{k,max}] \\ 0 & \text{otherwise} \end{cases}.$$

In an optimization scenario we can use Equation 3.11 to determine the M locally optimal prices for the SSO's revenue, corresponding to the M decision regions where at least one MNO is offloading. We then compile these values and find the globally optimal price for the SSO to set using the relationship in Equation 3.6.

3.2.3 Full SSO Utility

Now that we understand the extent of the problem, we can move forward and find the price that the SSO should charge to maximize its utility. Once again, the SSO's utility is different from its revenue, as it contains a term that acts as a hedge against the Quality of Service (QoS) burden that the influx of offloaded MNO users places on its network. This is an important consideration for the SSO as it is also concerned with the satisfaction of its own subscriber base.

To begin, we use the same relationship for the number of users that the MNO will offload in response to a particular offloading price, since we are still using the same piecewise-quadratic relationship for the MNO's cost function. In a particular price region k , the sum of offloaded users from all k MNOs in response to the price χ will be:

$$n_{off}^k = B_k - \chi D_k.$$

When we plug this relationship into the SSO's utility function listed in Equation 3.2, we find that the SSO's utility curve for a particular decision region k is:

$$U_{ss0,k}(\chi) = -D_k\chi^2 + B_k\chi + d_1n_{ss0}\log\left(\frac{E}{n_{ss0} + B_k - \chi D_k}\right).$$

This optimization becomes very similar to the single MNO case presented in [1]. Taking the derivative of $U_{ss0,k}$, we proceed:

$$U'_{ss0,k}(\chi) = B_k - 2D_k\chi + \left(\frac{d_1n_{ss0}D_k}{n_{ss0} + B_k - D_k\chi}\right) = 0$$

$$\chi^2(2D_k^2) + \chi(-3B_kD_k - 2D_kn_{ss0}) + (n_{ss0}B_k + B_k^2 + d_1n_{ss0}D_k) = 0.$$

Now that the LHS expression is in quadratic form we can easily find the solution $\chi_{U_k}^*$ that optimizes the extension of the curve found in Region k of the SSO's decision space:

$$\chi_{U_k}^* = \frac{1}{4D_k^2} \left[(3B_kD_k + 2n_{ss0}D_k) - \sqrt{(3B_kD_k + 2n_{ss0}D_k)^2 - 8D_k^2(n_{ss0}B_k + B_k^2 + d_1n_{ss0}D_k)} \right]. \quad (3.12)$$

In the same fashion as the revenue-only case, this value $\chi_{U_k}^*$ only maximizes the continuation of the portion of the SSO's utility curve seen in region k . We once again specify the value $\chi_{k,U}^*$ that actually maximizes the SSO's utility within region k :

$$\chi_{k,U}^* = \theta \cdot \chi_{U_k}^* + (1 - \theta) \left[\arg \max_{\chi \in \{\chi_{k,min}, \chi_{k,max}\}} U_{ss0}(\chi) \right]. \quad (3.13)$$

Continuing to follow the method set forth in the revenue-only case, we compile all these values of $\chi_{k,U}^*$ and use Equation 3.6 to find the globally optimal price that maximizes the SSO's utility.

Table 3.1: Single MNO Simulation Parameters

n_T	α	C	d_1	E
200	0.001	100	0.08	1000

3.3 Direct Implications of the Model

In order to understand how the QoS term in the SSO's utility function affects the SSO's behavior, we first consider the scenario in which only one MNO has an offloading agreement with the SSO (that is, $M = 1$). We independently performed two simulations for this scenario using the parameters in Table 3.1.

The first simulation assumed a constant $n_{ss0} = 30$ and varied n_{mno} between 200 and 300 to observe how this affected the number of users offloaded at the SSO's optimal price, both for the simplified SSO utility G and the full utility U_{ss0} . We denote n_G^* as the number of users offloaded at the optimal price when the SSO is trying to simply maximize its revenue, and n_U^* as the number of users offloaded when the SSO is trying to maximize its utility. The second experiment assumed a constant $n_{mno} = 300$ and instead varied n_{ss0} with the intent of observing its relationship to n_U^* .

These simulations offer insight into the applicability of our model. First of all, Figure 3.5 shows us that χ_U^* is consistently higher than χ_G^* . A direct result of this is that n_U^* is consistently lower than n_G^* , as seen in Figure 3.4. In fact, for lower values of n_{mno} , the SSO's utility is maximized by accepting zero offloaded users, and therefore the price it sets for these network conditions is arbitrarily high (in order to ensure that the MNO does not offload any

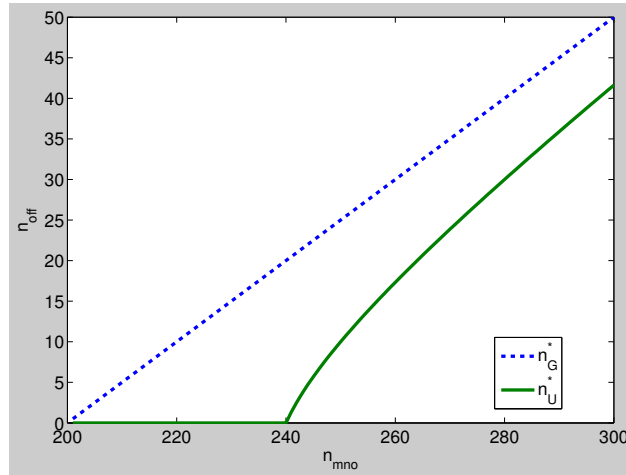


Figure 3.4: Equilibrium number of MNO users offloaded from the cellular network to the SSO's WLAN, as the total number of active MNO users (n_{mno}) varies. n_G^* is the number of offloaded users when the SSO is acting on its revenue function alone, and n_U^* is the number when the SSO acts on its full utility function.

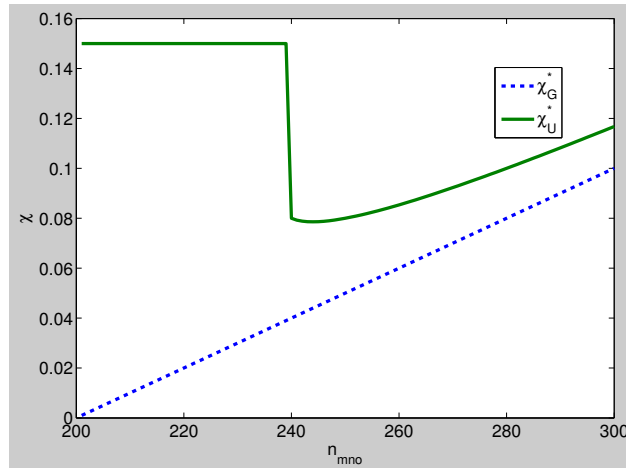


Figure 3.5: Equilibrium offloading price per user, set by the SSO, as the total number of active MNO users (n_{mno}) varies. χ_G^* is the price when the SSO is acting on its revenue alone, and χ_U^* is the price when the SSO acts on its full utility function.

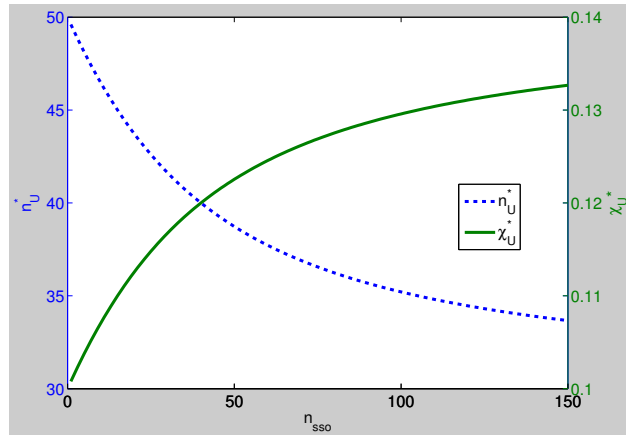


Figure 3.6: Equilibrium number of MNO users offloaded n_U^* and price χ_U^* per offloaded user, as the number of native SSO customers (n_{sso}) varies.

users). The reason for this is that the SSO has a constant positive utility when zero users are offloaded due to the *log* term in its utility function. As soon as n_{off} begins to increase, the value of its utility accounted for by this *log* term begins to decrease, since n_{off} is in the denominator of the *log* argument. Under the parameters of this particular simulation, when the MNO has only one active subscriber past its threshold n_T , if the SSO were to set the offloading price so low that this one user is offloaded, the increase in utility they would see from offloading revenue ($1 * \chi$) would be less than the decrease in utility they would see from the *log* term. This truth holds if the excess number of MNO users (that is, $n_{mno} - n_T$) is two, or three, or four, or higher, up to a certain transitional number. At this transitional number of MNO subscribers, there is a price the SSO can charge that would elicit a number of MNO subscribers to be offloaded such that the revenue gained by this action would be greater than the decrease in the log term of its utility function. From this point on, as the

MNO gains more subscribers past its threshold, there is never again a scenario where the SSO benefits by not having any users offloaded onto its network.

In the same vein, we find that as the number of native users n_{ss0} on the SSO's network increases, n_U^* decreases (seen in Figure 3.6). When n_{ss0} is large, the *log* term in the utility forces the SSO to raise its offloading price and accept fewer additional users onto its network than it otherwise would. This is in keeping with the idea that preserving quality of service (by limiting the number of users) is an important incentive.

An interesting effect visible in Figure 3.4 that is not immediately intuitive is that the gap between n_G^* and n_U^* diminishes as n_{mno} increases. This can be explained by considering revenue and QoS as a direct tradeoff. As n_{mno} grows, the MNO is willing to pay enough money to offload customers that the SSO feels it is adequately compensated for the QoS strain that it places on its network. In the snapshot that we observe this appears to be detrimental to the user experience of the SSO's prior users. However, the SSO's decision to sacrifice immediate QoS could be justified by its ability to upgrade its network with the additional revenue earned from the MNO. In the long term, the SSO could re-invest its earnings to appease its native customers.

3.4 Generalized Inferences

3.4.1 MNO Characteristics That Affect Offloading

In this work we study piecewise cost functions whose entire portion after the single threshold point, $n_{T,i}$, is strictly convex and strictly increasing. For the cost functions that we are considering in this thesis (quadratic, as well as the general polynomial form and exponential in the next chapter) the derivative of the convex portion will also be convex and strictly increasing. The inverse function of this derivative will be concave and strictly increasing. The function that we obtain in Equation 3.4 for the number of users that an MNO will offload is the negative of the concave and strictly increasing inverse function, and therefore it is a convex and strictly decreasing function. Therefore, if an MNO has any cost function whose increase after the threshold $n_{T,i}$ is strictly convex and strictly increasing, the number of users it chooses to offload will necessarily reach zero as the offloading price rises. Our aim in this section is to use this observation to learn about the characteristics that affect offloading.

The particular cost function that we explore in this paper has a quadratic increase and therefore a linear derivative. For an individual MNO, the number of users it will offload for a given price χ is provided in Equation 3.8:

$$n_{off,i} = b_i - \frac{\chi}{2\alpha_i}.$$

The χ -intercept of this function for MNO i , at which point $n_{off,i}$ hits zero and remains at zero for all higher values of χ , can easily be determined:

$$\chi_{0,i} = 2\alpha_i b_i. \quad (3.14)$$

We already know from Equation 3.8 that the two cost function factors that determine the $n_{off,i}(\chi)$ relationship are α_i , which influences the “steepness” of the convex increase after the threshold, and the value $b_i = n_{mno,i} - n_{T,i}$, which is the number of active MNO subscribers in excess of the threshold $n_{T,i}$. The relationship in Equation 3.14, however, tells us exactly how these cost function parameters affect *when* an MNO will begin to offload. That is, if the SSO were to initially set its offloading price very high so that no MNO wanted to offload, but then continually lower the price, it would observe that a particular MNO i will begin to offload after the price falls below its respective χ -intercept, $\chi_{0,i} = 2\alpha_i b_i$.

This is more easily understood in the context of Figure 3.2, where we see the offloading relationships of three separate MNOs. These MNOs have cost functions of the same quadratic form but with different characteristics. The product of $2\alpha_1$ and b_1 , which is the χ -intercept of MNO 1, is higher than the corresponding products of $2\alpha_i$ and b_i for MNOs 2 and 3. Therefore, MNO 1 is the first to offload as the price χ falls from high values.

3.4.2 Balance of Offloading Among Several MNOs

Generally, we assume that lower initial infrastructure costs will yield a higher value for both α_i and b_i , given the same network activity. To consider this intuitively, suppose an MNO only

invests in macrocell infrastructure when it first deploys its 4G network. The total investment cost will be low but the network will be severely capacity-limited. The capacity limitation would likely yield a lower value for $n_{T,i}$ in our model, which will in turn cause a greater value for b_i , which is one of the key indicators for the $n_{off,i}$ relationship. In addition to a lower threshold $n_{T,i}$, this network, when congested, will also more quickly cause dissatisfaction among customers and show signs of needing significant upgrades.

These factors, when incorporated into our model, lend themselves to a faster-rising cost function, which would be accounted for by a higher value of α_i . In such a scenario, we might observe the relationship seen in Figure 3.7, where the SSO maximizes its utility by setting the offloading price high enough that the only MNO that will choose to offload is the one that is severely capacity-constrained. In this sense, the other MNOs are *prevented* from offloading, and must bear the load of a surplus of users on their native networks.

Now consider an alternate scenario where each MNO makes a balanced initial investment in their cellular network, though still having minor variations in strategy among the group. When applied to our model, this situation would likely produce a lower variation among the several MNOs for both the cost function threshold $n_{T,i}$ and the steepness of the convex increase. Rather than yielding a single MNO that is desperate to offload even at a high price, this tighter grouping of cost function parameters would balance among the several MNOs the need to offload (assuming their active subscriber bases are comparable in size), and would therefore result in a lower offloading price that allows each MNO to take part in the offloading benefit. This scenario is depicted in Figure 3.8.

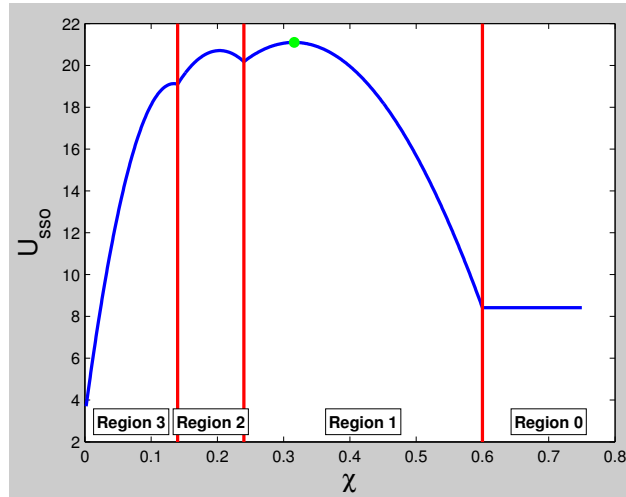


Figure 3.7: The globally optimal price χ_U^* falls in Region 1, where only a single MNO will offload.

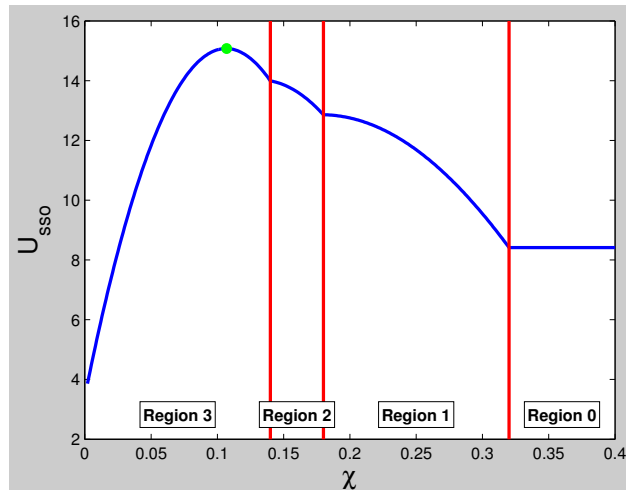


Figure 3.8: The globally optimal price χ_U^* falls in Region 3.

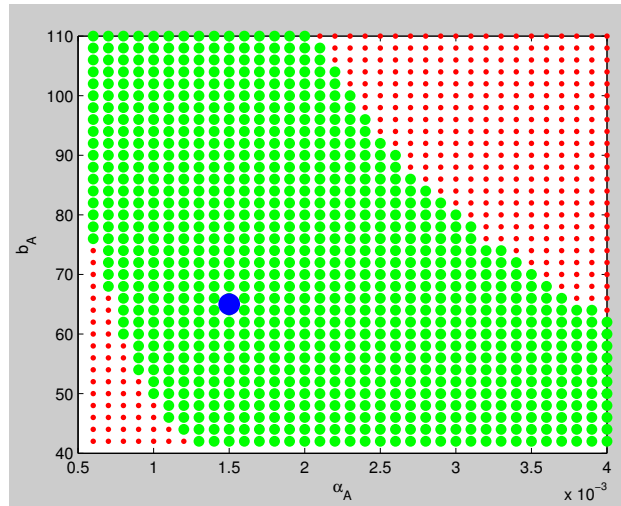


Figure 3.9: The effect of the variation of MNO parameters on the offloading scenario.

In order to understand this result more fully, consider Figure 3.9, which is the result of a simulation of two MNOs. We varied the cost function parameters of one MNO (which we'll call MNO A) while keeping those of the other MNO (MNO B) constant in order to observe at what point one of the two MNOs would “dominate” the offloading scenario. The large blue point indicates the fixed parameters of MNO B ($\alpha_B = 0.0015$, $b_B = 65$) and every other point indicates a simulated combination of the parameters of MNO A. The mid-sized green points represent the parameter combinations which prompted the SSO to set a price that allows both MNOs to offload active subscribers, while the small red points represent parameter combinations in which the offloading price only enticed a single MNO to offload. The red points for low values of α and b indicate scenarios where MNO B dominated offloading, while the red points for high values of α and b indicate scenarios where MNO A dominated offloading. As we can see, a higher value of one or both parameters can change the SSO's decision.

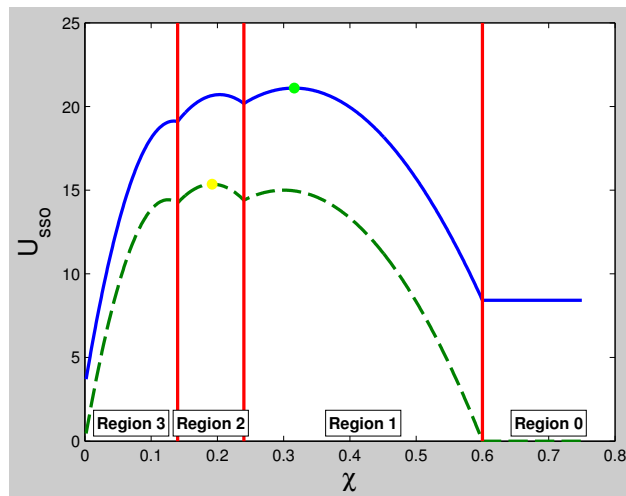


Figure 3.10: Maximizing the SSO's revenue function dictates a price that entices two MNOs to offload, whereas the full utility function dictates a price that would entice only one MNO to offload.

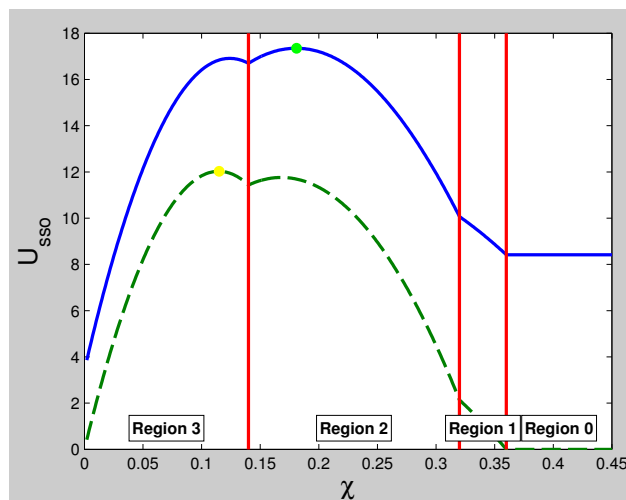


Figure 3.11: Maximizing the SSO's revenue function dictates a price that entices three MNOs to offload, whereas the full utility function dictates a price that would entice only two MNOs to offload.

In the context of this balance of offloading across the several MNOs, we examine further our previous observation about the SSO's utility function itself. In the case of a single MNO, the QoS term in the SSO's utility function causes the SSO to raise its price and require a greater number of active MNO subscribers to be present before the MNO is willing to offload. We observe an analogous phenomenon in the case of many MNOs. Consider Figure 3.10, which contains the same simulation parameters as Figure 3.7. The solid blue piecewise curve represents the SSO's utility, which, as the figure shows, is maximized by setting a price within region 1. The dashed green curve in the same figure represents the SSO's utility for the same simulation parameters, and we can see that this relationship is maximized by setting a price within region 2. This shows us that the QoS term in the SSO's utility function can fundamentally change the offloading scenario by causing the SSO to set a higher price that may prevent an MNO from offloading. The same effect is also visible in Figure 3.11, where the SSO's revenue is maximized by setting a price in region 3 and its utility is maximized by a price in region 2.

It should be noted that this effect is visible under very specific simulation parameters. In order to develop a clearer idea of how common this might be in practice, we set up a simulation with randomized MNO parameters. For an offloading scenario of three MNOs, the values of α_i varied between 0.0005 and 0.0035 and the values of b_i varied between 50 and 100. For each iteration of the simulation we paired the highest values of each parameter, the middle values of each parameter, and the lowest values of each parameter. This was to create the greatest possible variance in the χ -intercept values for each group of randomly generated

MNO parameters in order to represent a reasonably diverse group of cost functions. Out of 1000 iterations of the simulation, we found that 118 of them, or 11.8%, allowed for the utility-maximizing price to fall within a different decision region from the revenue-maximizing price. This simulation provides a clearer picture of the likelihood of an MNO in a real-world scenario being prevented from offloading due solely to the difference between the SSO's revenue function and utility function.

3.4.3 SSO with Limited Information

The entirety of this exploration thus far has been based on the assumption that the SSO has complete information. Optimization of the SSO's utility requires an explicit function to optimize, which is wholly based on prior knowledge of the MNO's cost function parameters. In a real-world offloading scenario, an MNO is unlikely to divulge its own cost function information. For this reason we conduct a brief exploration of a limited-information scenario.

In order to consider a scenario where the SSO has no knowledge of the MNO's cost function parameters, we designed a simulation where each MNO had random values for α_i and b_i within certain ranges. We still articulated the objective utility function for the SSO, but rather than using our result for the value that finds the optimal price, we instead employed a derivative-free optimization technique that used repeated evaluations of the SSO's utility at different prices (within a predetermined price range) until the optimal price was found within a certain error. This mimics the scenario in which the SSO can only learn

Table 3.2: Limited information simulation results (less restrictive accuracy requirements).

M	z_{avg}	z_{min}	z_{max}	Percent Accuracy
1	20.86	5	29	98.6
2	20.73	11	32	98.8
3	21.35	13	34	96.3
4	21.52	13	31	92.5
5	21.43	12	32	90.3

the behavior of the MNOs by methodically changing its price and observing the MNOs' response with each iteration. The optimization method used was a pattern search.

Our simulation separately considered scenarios with different values of M (number of MNOs who have an offloading agreement with the SSO). For each value of M , we ran 1000 test trials. Each test trial consisted of setting random values for α_i and b_i for each MNO, using the aforementioned optimization method, and recording the number of function evaluations (denoted by z) necessary before the optimal price was found. Over the 1000 test trials for each value of M , we recorded the average number of function evaluations to find the optimal price, the minimum number of function evaluations, and the maximum number of function evaluations. Tables 3.2 and 3.3 contain the results of the same simulation executed with slightly different parameters.

For the first execution of the simulation (results in Table 3.2) we used lax requirements on the optimization function. We considered an "accurate" solutions to be within 3% of

Table 3.3: Limited information simulation results (more restrictive accuracy requirements).

M	z_{avg}	z_{min}	z_{max}	Percent Accuracy
1	31.75	12	43	98.3
2	31.92	19	46	99.1
3	32.60	19	46	99.0
4	32.33	19	48	98.8
5	32.79	19	47	98.6

the exact answer calculated using the analytical methods previously explained in this chapter. Furthermore, the stopping criteria had low precision, which resulted in the percent of accurately determined solutions dropping when more MNOs are factored into the scenario.

For the second execution of the simulation we imposed both a greater accuracy requirement (solutions must be within 1% of the exact answer to be considered “accurate”). We also imposed more precise stopping criteria, which caused the percentage of accurate solutions to always be above 98% and not vary significantly for different numbers of MNOs, despite the greater accuracy requirement. The downside of this improved accuracy is that it generally requires a greater number of function evaluations.

An interesting observation we can make from these results is that the number of MNOs that have an offloading relationship with the SSO does not significantly affect the number of function evaluations it takes the SSO to arrive at the optimal price (though, with less precise stopping criteria, the accuracy of the solution *is* affected by the number of MNOs). Another

important observation is that the average number of iterations required is relatively low. One area of possible future work on this topic is developing a time-evolving model which centers around the fact that the SSO has limited information and must adaptively change its price over time in order to maximize its utility in response to a varying population of active MNO subscribers.

3.4.4 Application

This insight into the relationship between initial network investment and the need to offload encourages further investigation. With more precise information, one can determine these relationships more specifically. An MNO could use this model along with its own cost structure for a more accurate cost-benefit analysis of the investment into deploying a new technology. As an example, suppose an MNO targets a timespan of x years between deploying an LTE network and a next-generation LTE-advanced network. With the availability of WLAN offloading, this MNO may not need to deploy an extremely robust LTE network. It is possible that the MNO would spend less money over the targeted x years by deploying a capacity-limited LTE network and expecting to continuously offload users in dense areas than it would by initially deploying a much more robust LTE network and not offload at all.

This topic has in fact been considered in economic research. Real options theory, which takes into account investment uncertainty, has been used to address how a mobile operator should plan its next network investment [51] [52]. In particular, the authors of [51] discuss

operators of 2.5G networks who had the option of deferring investment in 3G networks for several more years or immediately investing in WLAN infrastructure to supplement their current network. In the same vein, MNOs currently need to make decisions regarding their 4G infrastructure investments. They can make large initial investments and rely on a robust cellular network or they can make limited initial investments and incur higher operating costs by buying WLAN access from other providers.

3.5 Conclusion

In this chapter we have put forth the bulk of the offloading framework addressed in this thesis. The SSO charges the MNO a price per offloaded user, and the MNO accordingly makes an offloading decision with its own cost function in mind. Though we have found some significant results, including the dynamic of allowing several MNOs to engage in an offloading relationship with an SSO, we still have not addressed in full the effect of the specific type of MNO cost function chosen. In the next section we employ the same model with different forms for the MNO cost function in order to see which results are universal and which will change.

Chapter 4

Different Cost Functions and Further Exploration

In this chapter we analyze certain extensions of our model and consider their applicability in the real world. In the first two sections we observe what the offloading relationship looks like when the MNO has a cost function with a different form from the quadratic function we have been using. We consider a polynomial cost function of a higher order than two as well as an exponential cost function, and observe how the results compare to those of the previous chapter. (Appendix A contains the derivation for most results in Sections 4.1 and 4.2.) Additionally, we perform a case study on a real-world metropolitan WLAN deployment

Section 4.3 is based on work included in [1], ©2014 IEEE.

in order to see how many active MNO subscribers would actually be “offloadable” (within range of Wi-Fi).

Finally, we take our base offloading model further by proposing an outline of a time-evolving model. The section that addresses this is not intended to provide a full exploration of the topic, but rather to present a limited analysis of a time-evolving model as an indicator of possible future work on this subject. The framework presented is based on representing the population of active MNO subscribers with a birth-death process.

4.1 General Polynomial Form Cost Function

We begin this exploration by examining scenarios where the MNOs have a polynomial cost function of order greater than two. Specifically, we examine a cost function of the form:

$$\psi_i(n) = \begin{cases} C_i & n \leq n_{T,i} \\ \alpha_i(n - n_{T,i})^{\beta_i} + C_i & n > n_{T,i} \end{cases}. \quad (4.1)$$

4.1.1 Number of Users Offloaded

Recall Equation 3.4 where we presented the relationship between n_{off} and χ for any cost function. We use this relationship and follow the same procedure outlined in Section 3.2.1:

$$\begin{aligned}\chi &= \psi'_i(n_i) \\ &= \alpha_i \beta_i (n_i - n_{T,i})^{\beta_i - 1} \\ n_i &= \psi'^{-1}_i(\chi) \\ &= \left(\frac{\chi}{\alpha_i \beta_i} \right)^{\frac{1}{\beta_i - 1}} + n_{T,i}.\end{aligned}$$

Therefore, for a particular MNO i ,

$$\begin{aligned}n_{off,i} &= b_i - \left(\frac{\chi}{\alpha_i \beta_i} \right)^{\frac{1}{\beta_i - 1}} \\ n_{off,i} &= b_i - \underbrace{\left(\frac{1}{\alpha_i \beta_i} \right)^{\frac{1}{\beta_i - 1}}}_{z_i} \chi^{\frac{1}{\beta_i - 1}} \\ n_{off,i} &= b_i - z_i \chi^{\frac{1}{\beta_i - 1}},\end{aligned}\tag{4.2}$$

where $b_i = n_{mno,i} - n_{T,i}$.

Proceeding forward, we only consider the scenario where the values of β_i are the same for each MNO (and we can simply call it β). We now determine the total number of users

offloaded onto the SSO's WLAN by all MNOs that are active in a given decision region k :

$$\begin{aligned}
 n_{off}^k &= \sum_{i=1}^k \left(b_i - z_i \chi^{\frac{1}{\beta-1}} \right) \\
 &= \underbrace{\sum_{i=1}^k b_i}_{B_k} - \underbrace{\sum_{i=1}^k \left(\frac{1}{\alpha_i \beta} \right)^{\frac{1}{\beta-1}}}_{Z_k} \chi^{\frac{1}{\beta-1}} \\
 &= B_k - Z_k \chi^{\frac{1}{\beta-1}}.
 \end{aligned}$$

4.1.2 Optimizing Prices

As with the exploration in the previous chapter, we begin by finding the revenue-optimizing price. Recall that the SSO's revenue, G , can be expressed by:

$$G = n_{off} \chi.$$

For a particular decision region k (which encompasses the price range that would entice exactly k MNOs to offload), we find that the price $\chi_{G_k}^*$ that optimizes the revenue curve seen in that region can be expressed by:

$$\chi_{G_k}^* = \left(\frac{B_k}{Z_k \left(\frac{1}{\beta-1} + 1 \right)} \right)^{\beta-1}. \quad (4.3)$$

Bear in mind that this price only maximizes the curve seen in region k of the SSO's decision space, just as with the scenario that we explored in Chapter 3. If $\chi_{G_k}^*$ does not fall within region k , the optimizing price for region k , represented by χ^*k, G , will be the boundary value of the region that produces the greater value for G (refer to Section 3.2.2 for a full explanation).

Due to the increased complexity of the cost function, we could not obtain a separated expression for the price $\chi_{U_k}^*$ that optimizes the curve of the SSO's full utility seen in region k . Despite this limitation, we are still able to find the optimal price in simulation using an exhaustive search method.

4.1.3 Comparison of Results With Quadratic Cost Function

A few key results of the general polynomial cost function remained the same as with the quadratic cost function:

- The QoS term in the SSO's cost function causes the SSO to set a consistently higher offloading price than it would if it were simply maximizing its revenue, both for the single MNO case and the multiple MNO case.
- When multiple MNOs have an offloading relationship with the SSO there are distinct decision regions for the SSO's price choice, and the number of MNOs able to offload for a particular scenario can vary depending on how the MNOs' cost function parameters are balanced.
- The complexity of determining the optimal offloading price when the SSO has limited information remains fairly low.

One significant difference between higher order polynomial cost functions and quadratic ones is the way in which the cost function parameters affect the n_{off} curve. If we set Equation

4.2 equal to zero, we find that the χ -intercept for a particular MNO's n_{off} can be expressed as:

$$\chi_{0,i} = \alpha_i \beta_i b_i^{\beta_i - 1}. \quad (4.4)$$

In the case of the quadratic cost function, the χ -intercept is directly proportional to α_i and b_i (see Equation 3.14). However, with a higher order polynomial cost function, the χ -intercept is proportional to a higher power of b_i .

4.2 Exponential Cost Function

Perhaps a more significant difference from the primary exploration is a piecewise-exponential cost function. Specifically, we examine a cost function of the form:

$$\psi_i(n) = \begin{cases} C_i & n \leq n_{T,i} \\ \alpha_i e^{\beta_i(n-n_{T,i})} + C_i & n > n_{T,i} \end{cases}. \quad (4.5)$$

4.2.1 Number of Users Offloaded

Recall Equation 3.4 where we presented the relationship between n_{off} and χ for any cost function. We use this relationship and follow the same procedure outlined in Section 3.2.1:

$$\begin{aligned}\chi &= \psi'_i(n_i) \\ &= \alpha_i \beta_i e^{\beta_i(n_i - n_{T,i})} \\ n_i &= \psi'^{-1}_i(\chi) \\ &= \frac{1}{\beta_i} \log\left(\frac{\chi}{\alpha_i \beta_i}\right) + n_{T,i}.\end{aligned}$$

Therefore, for a particular MNO i ,

$$\begin{aligned}n_{off,i} &= b_i - \frac{1}{\beta_i} \log\left(\frac{\chi}{\alpha_i \beta_i}\right) \\ n_{off,i} &= \underbrace{b_i + \frac{1}{\beta_i} \log(\alpha_i \beta_i)}_{y_i} - \underbrace{\frac{1}{\beta_i} \log(\chi)}_{z_i} \\ n_{off,i} &= y_i - z_i \log(\chi).\end{aligned}\tag{4.6}$$

where $b_i = n_{mno,i} - n_{T,i}$. Now we can determine the total number of users offloaded onto the SSO's WLAN by all MNOs that are active in a given decision region k :

$$\begin{aligned}n_{off}^k &= \sum_{i=1}^k (y_i - z_i \log(\chi)) \\ &= \underbrace{\sum_{i=1}^k \left(b_i + \frac{1}{\beta_i} \log(\alpha_i \beta_i)\right)}_{Y_k} - \underbrace{\sum_{i=1}^k \frac{1}{\beta_i} \log(\chi)}_{Z_k} \\ &= Y_k - Z_k \log(\chi).\end{aligned}$$

4.2.2 Optimizing Prices

As with the exploration in the previous chapter, we begin by finding the revenue-optimizing price. Recall that the SSO's revenue, G , can be expressed by:

$$G = n_{off} \chi.$$

For a particular decision region k , we find that the price $\chi_{G_k}^*$ that optimizes the revenue curve seen in that region can be expressed by:

$$\chi_{G_k}^* = e^{\frac{1}{Z_k}(Y_k - Z_k)}. \quad (4.7)$$

Once again, due to the increased complexity of the cost function, we could not obtain a separated expression for the price $\chi_{U_k}^*$ that optimizes the curve of the SSO's full utility seen in region k . Despite this limitation, we are still able to find the optimal price in simulation using an exhaustive search method.

4.2.3 Comparison of Results With Quadratic Cost Function

We used the same metrics to compare the results of the exponential cost function to those of the quadratic cost function as we did in the prior case of the higher-order polynomial cost function. Once again, the key results of the model were quite similar: The utility-maximizing price was consistently higher than the revenue-maximizing price, there were distinct regions for the SSO's price decision, and the complexity of learning the optimal price with limited information was relatively low.

Where we saw a distinct difference between the results of the exponential and polynomial cases was the way in which the cost function parameters affected the χ -intercept of the $n_{off,i}$ curves of the several MNOs. Using the same methods as before, we found that the χ -intercept can be expressed by:

$$\chi_{0,i} = \alpha_i \beta_i e^{\beta_i b_i}. \quad (4.8)$$

Once again, we see that the value of b_i affects the χ -intercept in an exponential manner. We performed a simulation with three MNOs where the values of α_i and b_i were randomly chosen within a certain range, while β_i remained the same (on the order of 10^{-1}) for each cost function through each iteration. Unlike the simulation described in Section 3.4.2, however, we did not match the greatest values of each parameter with each other. This complete randomness was to create a strong balance between the three MNOs of the combination of their parameters. The goal of the simulation was to see what proportion of the time the highest χ -intercept corresponded to the highest value of b_i , in order to measure the strength of influence of this parameter. Out of 1000 iterations of this simulation, there were 815 instances (or 81.5%) where the MNO with the largest value of b_i also had the largest χ -intercept. This is a strong indication that an MNO with a large subscriber base is far more likely to offload users at a higher price, even if its cost function is not as sharply increasing as those of other MNOs. The MNO with the highest χ -intercept is guaranteed the opportunity to offload (assuming the price is low enough that at least one MNO can offload), and is more likely to influence the offloading opportunity for other MNOs. It is not necessarily a

Table 4.1: Case Study Parameters

ρ_{in}	ρ_{out}
0.19	0.36

benefit to negatively impact the opportunity of other MNOs, but it is certainly a benefit for a particular MNO to not be negatively impacted.

4.3 Case Study of Offloadable Users

We applied our model in a simulation using information from a real municipal WLAN environment in order to understand how many users would actually be offloadable (within range of Wi-Fi). The city of Oulu, Finland operates a Wi-Fi network (called “panoulu”) in select areas throughout the city. We examined an area in downtown Oulu that is the busiest commercial area of the city (Figure 4.1). It is this type of region that is most likely to benefit from an offloading relationship like the one we are describing.

We have access to information on access point placement in the panoulu network throughout the city of Oulu. We use this information while varying certain MNO network parameters (including cost function parameters and average number of active subscribers) within a range that encompasses reasonable real-world conditions. This allows us to assess the performance

Panoulu AP location information accessible online at <http://www.panoulu.net/panoulu-wlan>



Figure 4.1: Area in downtown Oulu, Finland observed in this case study. We assessed the WLAN coverage within this area and determined the achievable offloading benefit when our model is applied. This figure is published in [1], ©2014 IEEE.

of a functioning WLAN deployment in conjunction with realistic MNO conditions in order to determine the applicability of our model.

The panoulu APs target certain indoor *and* outdoor areas. (The panoulu web site indicates that certain APs are designated for outdoor coverage.) We used this information as a basis for determining the total proportion of outdoor area covered within the observed zone, assuming a 50m outdoor propagation radius. The other APs listed were assumed to mostly cover indoor areas. If a sole AP was listed at a location we considered it to cover a circular indoor area with radius 20m. On the other hand, locations that listed multiple APs were considered to cover a certain fraction of the 100m by 100m square block area according to a rule that 12 APs would cover the full square block. Using these criteria we were able to determine the overall proportion of indoor area covered (ρ_{in}) as well as the proportion of outdoor area covered (ρ_{out}) within the section of the city observed for this study.

We classify a certain fraction c of the n_{mno} subscribers as indoor users and the rest as outdoor users. The total number of MNO subscribers within range of WLAN coverage, W , can then be expressed by the following linear function of c :

$$W = c\rho_{in}n_{mno} + (1 - c)\rho_{out}n_{mno}. \quad (4.9)$$

Cisco reports that nearly 80 percent of mobile data users are “either indoor or nomadic, rather than truly mobile” [53]. In light of this, we considered c to range from 0.5 to 0.9. For “worst case scenario” comparison purposes, we chose to compute the *minimum* number of offloadable users for each value of n_{mno} between 200 and 300 (in accordance with the

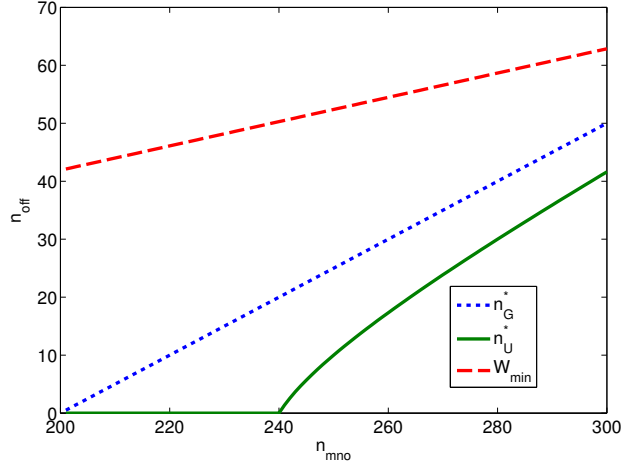


Figure 4.2: W_{min} , the minimum number of offloadable users, is greater than the number of users the MNO would offload for both the SSO’s utility-maximizing case and revenue-maximizing case. This figure is published in [1], ©2014 IEEE.

single-MNO simulation outlined in Section 3.3). Since $\rho_{in} < \rho_{out}$ (as seen in Table 4.1), Equation 4.9 is a decreasing function of c . Therefore the minimum value of W for any n_{mno} corresponds to the maximum value of c . For this reason we set $c = 0.9$, the highest value of our acceptable range.

Figure 4.2 shows that the minimum number of offloadable MNO subscribers is never lower than the number of users that the MNO would offload in the single-MNO simulation outlined in Section 3.3, both for the SSO’s utility-maximizing case and revenue-maximizing case. This means that, for the WLAN deployment currently in place in Oulu and with our simulated MNO network conditions, neither the MNO nor the SSO will find that their actions are limited by the number of users that are within range of the WLAN. We believe that this supports the applicability of our model in scenarios outside of Oulu as well.

4.4 Outline of a Time-Evolving Offloading Model

Throughout the entirety of this thesis, we have only analyzed a “snapshot” in time with a fixed number of users. This section aims to introduce a time-varying model. For simplicity of analysis, we only consider the scenario where a single MNO has an offloading relationship with the SSO and the SSO is maximizing its revenue. The key aspects changed from the original model are the following:

- The offloading price is now a price per user *per unit time*. The SSO is allowed to periodically change the price to account for the MNO’s usage. In our model, this would likely represent changing the price once every half-hour or so based on real-time usage, but it could also represent changing the price once every month based on average usage.
- Though the SSO can only change its price every so often, the MNO can choose to offload or not offload users much more frequently (corresponding with the unit time for which the offloading price is charged). Therefore, for a period of time the price is constant but the number of MNO users being offloaded is varying.
- The total number of active MNO subscribers $N(t)$ at any given time follows a birth-death process with exponential interarrival times. This includes both the users that the MNO chooses to service on its own network and those that it chooses to offload. The number of users the MNO wants to offload will vary with time corresponding with $N(t)$.

4.4.1 Detailed Changes to the Model

Let $N(t)$ follow a birth-death process path and be the total number of subscribers that an MNO has to service at a given time. Our initial analysis is concerned with this process when it's in equilibrium, which is when it has an equal probability of experience a birth or a death. We assume that at any given time, enough MNO users are within range of the SSO's WLAN access that the MNO will be able to offload as much as it wants.

The MNO's cost function, $\psi(N)$, represents the total cost to operate its native network over a time period of length h . This function accounts for the MNO's costs in the same way as the original model, and it includes the same threshold number of users n_T at which point the cost ceases to be constant and instead increases convexly. Once again, the key difference is that the total number of active MNO subscribers is changing over time, and therefore so is the number of offloaded users and the number of users serviced natively on the MNO's network. Formally, this cost function is expressed as:

$$\psi(N) = \begin{cases} C & N \leq n_T \\ \alpha(N - n_T)^2 + C & N > n_T \end{cases}. \quad (4.10)$$

We've defined the MNO's utility function in the same manner as the original model, except that it is now examined over a brief period of time h , which is the unit period of time for the SSO's offloading price:

$$U_{mno} = -h [\psi(N - N_{off}) + N_{off}\chi]. \quad (4.11)$$

The SSO's utility function simply considers the revenue it receives as a result of servicing offloaded users:

$$U_{sso} = \gamma N_{off} \chi, \quad (4.12)$$

where γ is the period of time over which the price must remain constant, and also the time duration for which the utility is observed. Bear in mind that χ is a price per unit time h , for which $\gamma \gg h$.

The same relationships presented in the original model regarding the total number of active subscribers and the number of offloaded users hold true here, except these numbers change over time:

$$N(t) - N_{off}(t) = N_b(t) \geq n_T. \quad (4.13)$$

In this equation $N_b(t)$ is the number of users that the MNO will service on its native network. Recall that the MNO will offload a user j if $\chi < \psi'_i(j)$. That is, the MNO will offload a customer only if the cost to do so is less than the cost to service the customer on the cellular network. Therefore when the MNO chooses the number of users it wants to offload in any given instant, we will approximate the relationship as $\chi = \psi'(N_b(t))$.

4.4.2 Preview of Analysis

This section is not intended to be a full analytical exploration. Rather, we simply set up the problem as a preview of potential future work in this area.

Before each time period of length γ when the SSO gets to readjust the offloading price that it charges (if it wants to), the strategy is to maximize its utility (defined in this model as just its revenue) for the subsequent time period. The analysis here would focus on optimizing the expected utility with respect to χ . In the case of an MNO population in equilibrium, this would likely yield an optimal price analogous to that of the static case presented in the prior chapter, as the time-average of a stable system often yields analytical results similar to the case of static analysis with a fixed mean.

There are times, however, when the MNO's network undergoes a period of sustained increase or decrease in users. For instance, in the several hours leading up to a major sporting event, the number of people in the vicinity of the stadium will likely grow. Alternatively, the hours between rush hour and bedtime on a weekday will likely see a prolonged decrease in MNO users to which the SSO must be able to react and change its price.

The SSO must have some strategy in place to detect these prolonged population changes. For example, perhaps the SSO could take advantage of control charts to monitor the number of users that are being offloaded. The purpose of control charts, according to [54], is “to detect out-of-control signals indicating the existence of special causes that affect the process stability by providing a visual indication of the behavior of critical quality variables.” Control charts have been the subject of a great deal of research, particularly in Statistics, and can be applied to theoretical stochastic processes as well as real-world systems. If the SSO recognizes a sustained increase or decrease pattern in the users that are being offloaded to its network, it can adjust its model such that the probabilities of transition between states

of the birth-death process that represents the MNO's population are unbalanced and there would be a net flow in one direction. In this scenario, the SSO would likely have to perform a less trivial analysis of their expected utility, taking into account the different amounts of revenue they would earn over the fixed-price period.

4.5 Conclusion

In this chapter we have explored the same offloading model presented in the previous chapter with some significant changes. First, we independently observed how two different cost function forms can have an effect on the offloading relationship, particularly the way in which a certain MNO's cost function parameters can affect the offloading landscape by providing a different χ -intercept for its n_{off} function. We also presented a case study that indicates what portion of active MNO subscribers would actually be within range of Wi-Fi in a real-world WLAN deployment. Finally, we presented a brief overview of a time-evolving offloading model using a birth-death process to represent the MNO's active subscriber base. This additional exploration provides a taste of what future work on this subject can entail. In the next chapter, we conclude this thesis by further describing potential future work on this topic and summarizing the key results and contributions of the thesis.

Chapter 5

Conclusion

In this thesis, we have considered the offloading relationship between MNOs and an SSO. We presented a model in which the SSO, in anticipation of an MNO's behavior, sets a strategic offloading price per user in order to maximize its utility.

One important result of the model itself is that the SSO's utility function imposes a tradeoff between the revenue it earns from offloaded users and the QoS burden on its network. The results in Section 3.3 show that the utility-maximizing price is consistently higher than the revenue-maximizing price. Furthermore, as the number of the SSO's own subscribers increases, we find that the SSO raises the offloading price in order to cause the MNO to offload fewer users.

More significant than the SSO's utility function itself, however, is what we learn about the nature of the offloading relationship. When multiple MNOs are present, the cost function

parameters and active subscriber population of one of them can affect the opportunity for other MNOs to offload users. If one MNO is disproportionately desperate to offload users at a higher price (caused by a high number of users and/or a steeply increasing cost function), the SSO may find that it can maximize its utility by setting a high price that prevents other MNOs from offloading. On the other hand, the network parameters are similar for the several MNOs, it's more likely that the offloading price will be in a range that allows more MNOs the opportunity to offload.

We found similar results even when we simulated different cost functions for the MNO. The key difference was the way in which the “ χ -intercept” (lowest price for which the MNO will offload zero users) was affected by the MNO's parameters. In the case of the exponential cost function, for instance, the excess of an MNO's active subscribers past its threshold is particularly significant as it affects the χ -intercept in an exponential manner. The reason this is significant is because the MNO with the highest χ -intercept is guaranteed the opportunity to offload (assuming the price is low enough that at least one MNO can offload), and is more likely to influence the offloading opportunities for other MNOs.

Though certain aspects of this investigation have been thoroughly explored, there is certainly room for future work. In this thesis, we have already suggested one such area in which the model would account for time-evolving network conditions. We believe a more thorough investigation in this area could provide a more robust understanding of how the offloading relationship would be conducted in practice. Another area in which this work can be expanded is AP-specific pricing. In a real-world scenario, an MNO's user base is not

necessarily evenly distributed. The fact that certain Wi-Fi APs may be in greater demand than others suggests that an MNO may be willing to pay more to offload certain users.

This seems to be a pivotal time for WLAN offloading in industry. With the move towards Hotspot 2.0 enabled devices, more mobile carriers are developing roaming agreements with Wi-Fi providers for the purpose of offloading traffic from congested cellular networks. This thesis provides a practical model for a cellular to Wi-Fi offloading agreement as well as insight into the nature of this relationship.

Bibliography

- [1] C. Patterson, A. MacKenzie, S. Glisic, B. Lorenzo, J. Roning, and L. DaSilva, “An economic model of subscriber offloading between mobile network operators and wlan operators,” in *12th International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt)*, May 2014, pp. 444–451.
- [2] Cisco, “Cisco visual networking index: Global mobile data traffic forecast update, 2012-2017,” 2013.
- [3] K. Johansson, A. Furuskar, P. Karlsson, and J. Zander, “Relation between base station characteristics and cost structure in cellular systems,” in *15th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, vol. 4, 2004, pp. 2627–2631 Vol.4.
- [4] J. Markendahl, Ö. Mäkitalo, J. Werding, and B. G. Mölleryd, “Business innovation strategies to reduce the revenue gap for wireless broadband services,” *Communications and Strategies*, 2009.

- [5] J. Björkdahl and E. Bohlin, “Competition policy and scenarios for European 3G markets,” *Communications & Strategies*, vol. 51, pp. 21–34, 2003.
- [6] N. Swartz, “The price of 3G,” *Connected Planet*, June 2001.
- [7] S. Grundberg, “Europe struggles to catch up with U.S. on 4G investment,” *The Wall Street Journal*, February 2014.
- [8] S. Liu and A. Striegel, “Casting doubts on the viability of Wi-Fi offloading,” in *Proceedings of the 2012 ACM SIGCOMM Workshop on Cellular networks: operations, challenges, and future design*, ser. CellNet ’12. New York, NY, USA: ACM, 2012, pp. 25–30. [Online]. Available: <http://doi.acm.org/10.1145/2342468.2342475>
- [9] L. Hu, C. Coletti, N. Huan, P. Mogensen, and J. Elling, “How much can WiFi offload? A large-scale dense-urban indoor deployment study,” in *IEEE 75th Vehicular Technology Conference (VTC Spring)*, 2012, pp. 1–6.
- [10] S. Frei, W. Fuhrmann, A. Rinkel, and B. Ghita, “Prospects for WLAN in the evolved packet core environment,” in *2012 5th International Conference on New Technologies, Mobility and Security (NTMS)*, 2012, pp. 1–5.
- [11] A. Balasubramanian, R. Mahajan, and A. Venkataramani, “Augmenting mobile 3G using WiFi,” in *Proceedings of the 8th international conference on Mobile systems, applications, and services*, ser. MobiSys ’10. New York, NY, USA: ACM, 2010, pp. 209–222. [Online]. Available: <http://doi.acm.org/10.1145/1814433.1814456>

- [12] B. Han, P. Hui, V. A. Kumar, M. V. Marathe, G. Pei, and A. Srinivasan, “Cellular traffic offloading through opportunistic communications: A case study,” in *Proceedings of the 5th ACM Workshop on Challenged Networks*, ser. CHANTS '10. New York, NY, USA: ACM, 2010, pp. 31–38. [Online]. Available: <http://doi.acm.org/10.1145/1859934.1859943>
- [13] G. Hiertz, D. Denteneer, L. Stibor, Y. Zang, X. Costa, and B. Walke, “The IEEE 802.11 universe,” *Communications Magazine, IEEE*, vol. 48, no. 1, pp. 62–70, January 2010.
- [14] Cisco, “802.11ac - The fifth generation of Wi-Fi, technical white paper,” 2012.
- [15] P. Deshpande, X. Hou, and S. R. Das, “Performance comparison of 3G and metro-scale WiFi for vehicular network access,” in *Proceedings of the 10th ACM SIGCOMM Conference on Internet Measurement*, ser. IMC '10. New York, NY, USA: ACM, 2010, pp. 301–307. [Online]. Available: <http://doi.acm.org/10.1145/1879141.1879180>
- [16] J. Andrews, H. Claussen, M. Dohler, S. Rangan, and M. Reed, “Femtocells: Past, present, and future,” *IEEE Journal on Selected Areas in Communications*, vol. 30, no. 3, pp. 497–508, April 2012.
- [17] S. Taylor, “The next generation of Wi-Fi debuts in Beijing,” Cisco Blog, 2013.
- [18] Trefis Team, “AT&T roams to \$38 thanks to iphone halo,” Forbes, 2010.
- [19] AT&T, “AT&T iphone customers to receive free access to nation’s largest Wi-Fi network,” AT&T Press Release, 2008.

- [20] —, “AT&T: 2.7 billion connections made in 2012; Wi-Fi network growth doubles,” AT&T Press Release, 2013.
- [21] K. O’Neill, “Boingo announces global Wi-Fi roaming agreement with AT&T,” Boingo Press Release, 2013.
- [22] —, “Boingo adds 75,000 hotspots in South Korea via SK Telecom,” Boingo Press Release, 2012.
- [23] —, “Boingo to provide global roaming and data offload solutions for NTT Docomo,” Boingo Press Release, 2012.
- [24] —, “Boingo adds new hotspots via China Telecom,” Boingo Press Release, 2010.
- [25] M. Dano, “Boingo inks Wi-Fi offload deal with Verizon, hopes for ‘significant revenue’ from Hotspot 2.0,” Fierce Wireless, 2014.
- [26] K. O’Neill, “NYC subway stations to get Wi-Fi via boingo and transit wireless agreement,” Boingo Press Release, 2012.
- [27] WBA, “Wireless Broadband Alliance and 13 operators to enable ‘NGH Live’ at Wi-Fi Global Congress Beijing,” WBA Press Release, 2013.
- [28] IEEE Computer Society, “Part 11: Wireless LAN medium access control (MAC) and physical layer (PHY) specifications, amendment 9: Interworking with external networks,” IEEE, Tech. Rep., 2011.

- [29] Cisco, “The future of hotspots: Making WiFi as secure and easy to use as cellular, white paper,” 2012.
- [30] IEEE Computer Society, “Part 11: Wireless LAN medium access control (MAC) and physical layer (PHY) specifications; amendment 2: Fast basic service set (BSS),” IEEE, Tech. Rep., 2008.
- [31] Z. Zhang, “Routing in intermittently connected mobile ad hoc networks and delay tolerant networks: Overview and challenges,” *IEEE Communications Surveys and Tutorials*, vol. 8, no. 1, pp. 24–37, 2006.
- [32] Y. Li, G. Su, P. Hui, D. Jin, L. Su, and L. Zeng, “Multiple mobile data offloading through delay tolerant networks,” in *Proceedings of the 6th ACM Workshop on Challenged Networks*, ser. CHANTS ’11. New York, NY, USA: ACM, 2011, pp. 43–48. [Online]. Available: <http://doi.acm.org/10.1145/2030652.2030665>
- [33] C. Mayer and O. Waldhorst, “Offloading infrastructure using delay tolerant networks and assurance of delivery,” in *2011 IFIP Wireless Days (WD)*, Oct 2011, pp. 1–7.
- [34] C. P. Mayer and O. P. Waldhorst, “Routing in hybrid delay tolerant networks,” *Computer Communications*, vol. 48, no. 0, pp. 44 – 55, 2014, opportunistic networks. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0140366414001042>
- [35] K. Lee, J. Lee, Y. Yi, I. Rhee, and S. Chong, “Mobile data offloading: How much can WiFi deliver?” *IEEE/ACM Transactions on Networking*, vol. 21, no. 2, pp. 536–550, April 2013.

- [36] N. Ristanovic, J.-Y. Le Boudec, A. Chaintreau, and V. Erramilli, “Energy efficient offloading of 3G networks,” in *2011 IEEE 8th International Conference on Mobile Adhoc and Sensor Systems (MASS)*, Oct 2011, pp. 202–211.
- [37] S. Sen, C. Joe-Wong, S. Ha, and M. Chiang, “Incentivizing time-shifting of data: A survey of time-dependent pricing for internet access,” *IEEE Communications Magazine*, vol. 50, no. 11, pp. 91–99, November 2012.
- [38] S. Yaiparoj, F. Harmantzis, and V. Gunasekaran, “On the economics of GPRS networks with Wi-Fi integration,” *European Journal of Operational Research*, vol. 187, no. 3, pp. 1459 – 1475, 2008. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0377221706008563>
- [39] X. Zhuo, W. Gao, G. Cao, and Y. Dai, “Win-coupon: An incentive framework for 3G traffic offloading,” in *2011 19th IEEE International Conference on Network Protocols (ICNP)*, Oct 2011, pp. 206–215.
- [40] J. Lee, Y. Yi, S. Chong, and Y. Jin, “Economics of WiFi offloading: Trading delay for cellular capacity,” in *2013 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, April 2013, pp. 357–362.
- [41] K. Zhu, E. Hossain, and D. Niyato, “Pricing, spectrum sharing, and service selection in two-tier small cell networks: A hierarchical dynamic game approach,” *IEEE Transactions on Mobile Computing*, vol. PP, no. 99, pp. 1–1, 2013.

- [42] D. Niyato and E. Hossain, “Wireless broadband access: WiMAX and beyond - integration of WiMAX and WiFi: Optimal pricing for bandwidth sharing,” *IEEE Communications Magazine*, vol. 45, no. 5, pp. 140–146, 2007.
- [43] G. Iosifidis, L. Gao, J. Huang, and L. Tassiulas, “An iterative double auction for mobile data offloading,” in *2013 11th International Symposium on Modeling Optimization in Mobile, Ad Hoc Wireless Networks (WiOpt)*, 2013, pp. 154–161.
- [44] W. Dong, S. Rallapalli, R. Jana, L. Qiu, K. Ramakrishnan, L. Razoumov, Y. Zhang, and T. W. Cho, “iDEAL: Incentivized dynamic cellular offloading via auctions,” in *2013 Proceedings of IEEE INFOCOM*, April 2013, pp. 755–763.
- [45] F. Zhang, W. Zhang, and Q. Ling, “Non-cooperative game for capacity offload,” *IEEE Transactions on Wireless Communications*, vol. 11, no. 4, pp. 1565–1575, April 2012.
- [46] L. Gao, G. Iosifidis, J. Huang, and L. Tassiulas, “Economics of mobile data offloading,” in *Proc. of IEEE SDP Workshop*, 2013.
- [47] H. Le Cadre, M. Bouhtou, and B. Tuffin, “A pricing model for a mobile network operator sharing limited resource with a mobile virtual network operator,” in *Network Economics for Next Generation Networks*, ser. Lecture Notes in Computer Science, P. Reichl, B. Stiller, and B. Tuffin, Eds. Springer Berlin Heidelberg, 2009, vol. 5539, pp. 24–35.

- [48] A. Eshghi, D. Haughton, and H. Topi, “Determinants of customer loyalty in the wireless telecommunications industry,” *Telecommunications Policy*, vol. 31, no. 2, pp. 93 – 106, 2007.
- [49] S. Aydin, G. Ozer, and O. Arasil, “Customer loyalty and the effect of switching costs as a moderator variable: A case in the turkish mobile phone market,” *Marketing Intelligence and Planning*, vol. 23, no. 1, pp. 89 – 103, 2005.
- [50] M. Lee and L. Cunningham, “A cost/benefit approach to understanding service loyalty,” *Journal of Services Marketing*, vol. 15, no. 2, pp. 113–130, 2001.
- [51] F. C. Harmantzis and V. P. Tanguturi, “Investment decisions in the wireless industry applying real options,” *Telecommunications Policy*, vol. 31, no. 2, pp. 107 – 123, 2007. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0308596106001066>
- [52] J. H. Alleman and E. M. Noam, *The new investment theory of real options and its implication for telecommunications economics*. Springer, 1999, vol. 34.
- [53] Cisco, “Cisco service provider WiFi: A platform for business innovation and revenue generation,” 2012.
- [54] W. Hachicha and A. Ghorbel, “A survey of control-chart pattern-recognition literature (1991 to 2010) based on a new conceptual classification scheme,” *Computers & Industrial Engineering*, vol. 63, no. 1, pp. 204 – 222, 2012. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S036083521200071X>

Appendix A

Derivations for Results in Sections 4.1 and 4.2

A.1 Revenue-optimizing Price for Polynomial Cost Function

Recall that:

$$B_k = \sum_{i=1}^k b_i,$$
$$Z_k = \sum_{i=1}^k \left(\frac{1}{\alpha_i \beta} \right)^{\frac{1}{\beta-1}}, \text{ and}$$
$$n_{off}^k = B_k - Z_k \chi^{\frac{1}{\beta-1}}$$

Now:

$$\begin{aligned} G(\chi) &= n_{\text{off}}\chi \\ &= \chi \left(B_k - Z_k \chi^{\frac{1}{\beta-1}} \right) \end{aligned}$$

$$\begin{aligned} G'(\chi) &= \chi \left(-\frac{Z_k}{\beta-1} \chi^{\left(\frac{1}{\beta-1}-1\right)} \right) + B_k - Z_k \chi^{\frac{1}{\beta-1}} = 0 \\ &-\frac{Z_k}{\beta-1} \chi^{\frac{1}{\beta-1}} + B_k - Z_k \chi^{\frac{1}{\beta-1}} = 0 \\ &\chi^{\frac{1}{\beta-1}} \left(-\frac{Z_k}{\beta-1} - Z_k \right) = -B_k \\ &\chi^{\frac{1}{\beta-1}} = \frac{B_k}{Z_k} \cdot \frac{1}{\frac{1}{\beta-1} + 1} \\ &\chi_{G_k}^* = \left(\frac{B_k}{Z_k \left(\frac{1}{\beta-1} + 1 \right)} \right)^{\beta-1} \end{aligned}$$

A.2 χ -intercept for Polynomial Cost Function

$$\begin{aligned} n_{\text{off},i} &= b_i - \chi^{\frac{1}{\beta_i-1}} \left(\frac{1}{\alpha_i \beta_i} \right)^{\frac{1}{\beta_i-1}} = 0 \\ &\chi^{\frac{1}{\beta_i-1}} (\alpha_i \beta_i)^{-\frac{1}{\beta_i-1}} = b_i \\ &\chi^{\frac{1}{\beta_i-1}} = b_i (\alpha_i \beta_i)^{\frac{1}{\beta_i-1}} \\ &\chi_{0,i} = \alpha_i \beta_i b_i^{\beta_i-1} \end{aligned}$$

A.3 Revenue-optimizing Price for Exponential Cost Function

Recall that:

$$Y_k = \sum_{i=1}^k \left(b_i + \frac{1}{\beta_i} \log(\alpha_i \beta_i) \right),$$

$$Z_k = \sum_{i=1}^k \frac{1}{\beta_i}, \text{ and}$$

$$n_{off}^k = Y_k - Z_k \log(\chi)$$

Now:

$$\begin{aligned} G(\chi) &= n_{off} \chi \\ &= \chi(Y_k - Z_k \log(\chi)) \end{aligned}$$

$$G'(\chi) = Y_k - Z_k(1 + \log(\chi)) = 0$$

$$Z_k \log(\chi) = Y_k - Z_k$$

$$e^{Z_k \log(\chi)} = e^{Y_k - Z_k}$$

$$\chi^{Z_k} = e^{Y_k - Z_k}$$

$$\chi_{G_k}^* = e^{\frac{1}{Z_k}(Y_k - Z_k)}$$

A.4 χ -intercept for Exponential Cost Function

$$\begin{aligned}n_{off,i} &= \left(b_i + \frac{1}{\beta_i} \log(\alpha_i \beta_i) \right) - \frac{1}{\beta_i} \log(\chi) = 0 \\ \log(\chi) &= \beta_i \left(b_i + \frac{1}{\beta_i} \log(\alpha_i \beta_i) \right) \\ e^{\log(\chi)} &= e^{\beta_i b_i} e^{\log(\alpha_i \beta_i)} \\ \chi_{0,i} &= \alpha_i \beta_i e^{\beta_i b_i}\end{aligned}$$