

Network Selection Strategy in Heterogeneous Wireless Networks

*Olga Ormond¹, Dr. Gabriel-Miro Muntean²,
Dr. John Murphy¹,

¹Department of Computer Science, University
College Dublin, Belfield, Dublin 4, Ireland,
Tel: +353-1-716-2943

E-mail: olga.ormond@ucd.ie

²School of Electronic Engineering, Dublin City
University, Dublin, Ireland

Abstract: Next generation communications will offer a wide range of services available to users anywhere at any time. How wireless users access those services depends on the current user context: location, characteristics of the available networks, user preferences, application requirements and terminal capabilities. The vision is that users will not be tied down to a long-term contract with one single operator and will instead be able to dynamically choose access provision on a per call basis. The evolving competitive marketplace will provide a choice of access networks in any given location, each offering different network technologies with varying characteristics to transport the user's communications application. This paper highlights the need for an access network selection decision strategy to aid users operating in this heterogeneous multi-network wireless environment. We propose a consumer surplus based algorithm that selects the best available network for transferring non real-time data, with user specified time constraints. When compared to an always cheapest strategy simulation results show significant performance gains in transfer completion time for the consumer surplus based strategy.

1. Introduction

There are a number of views on how future generations of mobile network systems will integrate and operate. Technologies are evolving with a focus on dynamic adaptation and interoperability. Internet services and more traditional mobile applications are merging onto the same devices, which are now being built as multi-homed multi-functioning wireless terminals with varying memory and processing capabilities. The regulators are endorsing a consumer friendly competitive marketplace. It is foreseen that service provider will emerge as a third party in this environment.

A Service Oriented Heterogeneous Wireless Network Environment (SOHWNE) is envisioned as a setting where network access will be decoupled from service provision. As depicted in Figure 1, it is anticipated that users and service providers will be free from long-term subscription contracts to any one Radio Access Network (RAN) operator and can instead dynamically select the most suitable transport offering from the available network providers for their application or service requirements [1]. For example Figure 1 shows how a user who requires a data service from service provider S3 and surveys the network availability at his/her current location discovers that access is available via either provider a network, or the

network offered by provider **b**. The user's decision of which network to utilise will then depend on his/her preferences and the condition of the networks in question.

This newly competitive marketplace will be driven by growing user demand for innovative online services available anywhere at any time. RAN providers will strive to generate revenue in this new SOHWNE market by charging for traffic transportation services. However, to earn revenue the cost-performance ratio for network access must satisfy the ordinary user, or the user will look elsewhere for network connectivity.

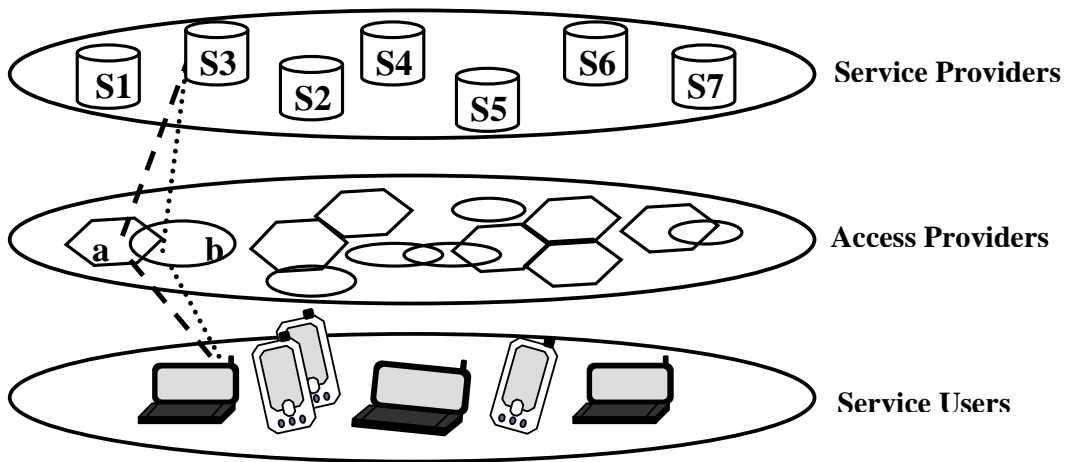


Figure 1 Service Oriented Heterogeneous Wireless Network Environment

In this context, users will strive to make the most of this consumer oriented environment. Naturally selfish, they will aim to get the best connection possible for their money. The difference between the actual price paid for communications and the utility or value of the data to the user is known in microeconomic terms as the Consumer Surplus (CS). Users aim to maximise savings by maximising positive CS. Next generation users will require a terminal and applications which are adaptable to the dynamic characteristics of the radio environment, they will also strive to own a device with intelligent inbuilt features to aid them to meet their particular communication preferences efficiently.

This paper introduces a novel RAN selection strategy for non real-time data users in SOHWNEs. The strategy is based on maximising users' consumer surplus while meeting the individual user specified performance requirements for transfer completion time. Available networks are short-listed according to their predicted ability to meet the data delay constraints. The optimal network is then chosen by maximising the user's CS and the data will be delivered over this network. In the next section we provide an overview of related research. This is followed by section 3 which details the proposed CS network selection strategy. Section 4 defines the testing scenarios and presents comparison results when CS and the always-cheapest network selection strategies were used respectively. This paper concludes with a discussion on the user satisfaction improvements when using the proposed CS-based strategy.

2. Related Work

A concentration of research in this area focuses on access network providers, and technology advancements to provide near seamless mobility in this multi-service multi-technology multi-user environment. Hot topics include fast and seamless handover, efficient billing systems, network revenue maximization and spectrum utilisation. Much work is also underway to optimize service provider offerings, adapting applications to individual user preferences, terminal capabilities and access technology.

The area of RAN selection policy design has received far less attention. Much of the work in this area investigates the access network selection problem as part of the seamless handover venture. McNair and Zhu outline in [2] some decision metrics and state the need for decision policy design in the context of vertical handover for a single mobile user running multiple communication sessions. This is also relevant to the initial network selection decision for any user with a choice in available RANs. In the case of a semi-stationary mobile terminal an initial intelligent network selection may save the user cost in avoiding handover costs.

Utility-based functions are commonly used to describe user preference rating relationship for a number of metrics. X. Wang et al [3] consider user preferences to be represented quantitatively through a utility function when comparing two congestion pricing schemes. In [4] Das et al consider users to choose a pricing plan based on their data delay considerations, described by a user utility function. They use their understanding of user behaviour to maximise network gains, however they do not consider SOHWNEs but rather look at efficient network resource management with the goal of maintaining a steady customer base (i.e. reducing customer churn) and maximizing wireless network operator revenue.

In [5], Ylitalo et al propose a solution for facilitating a user making a network interface selection decision. They focus mainly on a possible architecture for the end terminal, and mention the example of an Always Cheapest Network Selection strategy. In this paper we assume that mobile terminals will be equipped with an architecture that can facilitate user network selection decisions, and concentrate on what actual decision strategy to use. The user network selection decision strategy will be influenced by the pricing scheme employed in the available networks. Le Bodic et al [6] look at networks with an auction based pricing scheme and employ two different strategies based on user preference for low service charge, or user preference for networks with a good reputation. Many new pricing schemes are being proposed for RANs [7], the majority focus on network-centric benefit. In this paper we assume all available networks employ a fixed price per byte pricing scheme, as i-mode users are currently charged.

The focus of this paper is on the network selection decision problem and the goal is to propose a solution which brings benefits to non real-time data users with intelligent selection criteria for which available network to use. In this context Bircher and Braun [8] propose an agent-based architecture with a user agent decision function. Customers compare and select services with the best performance/price ratio, negotiate with providers for offered services and pre-reserve the resources for an agreed price. Details of the exact negotiation terms are not covered but they are based on differently weighted QoS parameters such as delay, bandwidth, packet loss, etc. The feasibility of reserving guaranteed bandwidth in unreliable radio environment is questionable.

Access discovery is part of a common three-phase handover procedure. In this phase the wireless terminal seeks information on available access networks. The

amount of information supplied should be minimal. Receiving current network condition information may be resource consuming and wasteful in the ever changing unreliable wireless environment. We use previous file delays experienced in a particular network to predict transfer completion time for that network. In [9] Ramani and Savage work on faster handoff between Wi-Fi networks. Their low cost technique for continuously tracking nearby base stations is a method which could readily provide a list of available networks, and their current status, for IEEE 802.11 networks.

3. CS (Consumer Surplus) Network Selection Strategy

This paper considers the wireless user scenario depicted in Figure 2. A user in a multiple-access network environment wants to access a non real-time data service. Specifically s/he requires transfer of a data file uplink from their radio terminal to a service provider's server in the wired network. For example the user may be uploading photos from a laptop, emailing a large document from their mobile phone, or uploading vending machine data from their PDA to a central server. These services may differ in characteristic requirements: amount of data to transfer, and delay provisions.

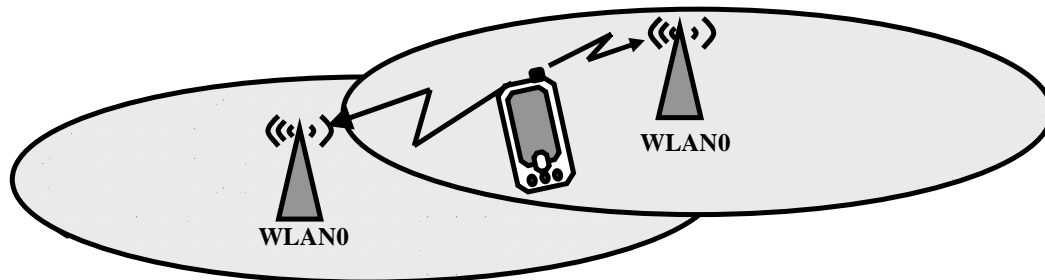


Figure 2 Choice of Access Network: WLAN0 or WLAN1

Dependent on their current location the user has a choice of several available networks. Each network in the system employs a fixed price per byte pricing scheme but charges different prices and is subject to different background traffic patterns. Naturally every user wants timely delivery of their data at the lowest price. While the strict time restrictions imposed by real-time traffic are not used here, it is assumed that every user has a patience limit and will only be willing to wait so long for completed transfer of their data before they become dissatisfied. The more delay a user experiences the less s/he may be willing to pay.

The user may employ different possible tactics to select the network which will maximise his/her satisfaction. For example a user may elect to always select one designated network regardless of its current characteristics, or to always minimise spending by choosing the cheapest network, or decide to continually opt for random network selection. However we believe that an intelligent selection based on user willingness to pay, file completion time constraints and estimated access network delivery time should be used instead. We propose a utility-based algorithm that accounts for user time constraints, estimates complete file delivery time (for each of the available access networks) and then selects the most promising network based on consumer surplus.

To accurately describe the trade-off between the user's time and budget limitations we introduce a typical piece-wise linear user utility function as shown in Figure 3. Other possible utility functions are discussed in the section on future work.

The typical utility function, as shown in (2), relates to the zone-based structure introduced in user studies [8]. Users' patience with file delivery completion time is classified into three zones: satisfaction zone, tolerance zone and frustration zone. These correspond to zones 1, 2 and 3 respectively in Figure 3.

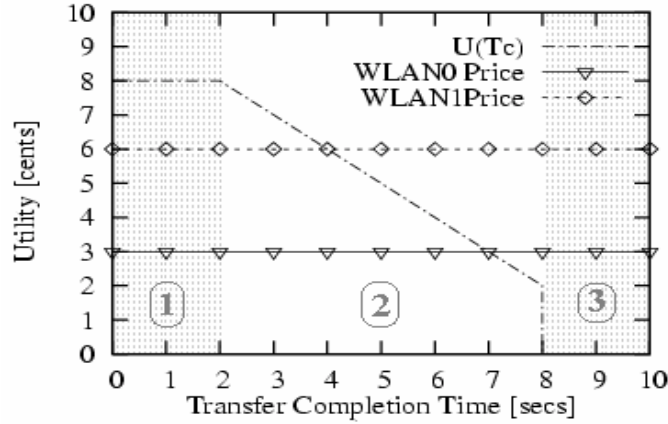


Figure 3. Typical User Utility function for 200 kByte payload

User's willingness-to-pay or utility U_i for a file i depends the required transfer completion time (T_c). T_c , in seconds, is related to size of the file and depends on the average end-to-end throughput rate according to (1). Where F_i is size of file i in bits, and r is average rate at which the data is transferred across the network. Timely arrival of the file anywhere within zone 1 is worth the highest value, U_{\max} . We denote completion time T_{c1} , as the threshold of the satisfaction zone. After waiting T_{c1} seconds the value of completed transfer linearly decreases with increase in delay, U_{var} , until the time exceeds the user defined maximum tolerated delay, T_{c2} . Any file arriving after this specified maximum completion time is worth 0 cent to the frustrated user.

$$T_c = F_i / r \quad (1)$$

$$U_i(T_c) = \begin{cases} U_{\max}, & T_c \leq T_{c1} \\ U_{\text{var}} - T_c, & T_{c1} < T_c \leq T_{c2} \\ 0, & T_c > T_{c2} \end{cases} \quad (2)$$

Table I shows the association between expected user delays and the measured bit-rates. It takes into account a large range of possible throughput rates available in existing and emerging RANs and a good range of likely file sizes. Depending on the RAN available the bit rates will vary, and what is presented here is only one set of rates. In this table any delay below 1 second is considered negligible, whereas delays greater than 30 seconds are breaching the end-users expectation for good service. Different users can have different expectations for file completion times, and therefore the shading in this table can vary.

TABLE 1
RATES, FILE SIZES AND CORRESPONDING DELAYS

File size	Filesize & Hdrs bytes	kbps 5000 secs	kbps 2000 secs	kbps 1000 secs	kbps 500 secs	kbps 200 secs	kbps 100 secs	kbps 50 secs	kbps 20 secs	kbps 10 secs
10	10660	0.02	0.04	0.09	0.17	0.43	0.85	1.71	4.26	8.53
20	21260	0.03	0.09	0.17	0.34	0.85	1.70	3.40	8.50	17.01
50	53060	0.08	0.21	0.42	0.85	2.12	4.24	8.49	21.22	42.45
80	84860	0.14	0.34	0.68	1.36	3.39	6.79	13.58	33.94	67.89
100	106060	0.17	0.42	0.85	1.70	4.24	8.48	16.97	42.42	84.85
200	212060	0.34	0.85	1.70	3.39	8.48	16.96	33.93	84.82	169.65

In order to provide good transport services this algorithm considers a bit-rate of 212 kbps as the threshold for user acceptability. The associated user utility function for an FTP application transferring a payload of 200 Kbyte is shown in Figure 3. The graph indicates how any completion time up to $T_{c1} = 2$ seconds is worth the highest price to the user, whereas files arriving after $T_{c2} = 8$ seconds are worth 0 cent. In between T_{c1} and T_{c2} the value linearly decreases with time.

$$CS_i = U_i(T_c) - C_i \quad (3)$$

The predicted CS_i can be calculated using (3) and cost functions for the file transfer and the user utility calculated using the predicted delay for transfer in each of the available networks. CS_i is consumer surplus in cent for transporting file i , $U_i(T_c)$ is the monetary value (in cent) that the user places on the transfer of file i in the given the transfer completion time (T_c) and C_i is the cost charged by the network, also in cent, for the completed file transfer.

The available access network options are considered with the aim to maximise the consumer surplus, subject to user-imposed time constraints related to the complete file transfer. In the case of each available RAN, the user terminal must predict the current offered rate and use those estimated rates, together with the provided utility function for the current application, to select the network that will give the greatest consumer surplus while meeting the transfer completion time deadline.

4. Testing and Results

4.1 Testing Setup and Scenarios

As described in Section 3, the scenario considered in this paper contains two overlapping WLANs. Each WLAN contains a number of associated background traffic generating mobile nodes. Both LANs are connected to a wired network which hosts the sink for all application data. We start with one intelligent user terminal located in the overlap area which has the option to choose either of the two RANs for FTP transfer of the data file. Later additional intelligent users are added to make the competing traffic more realistic and therefore the model and results more accurate. During successive tests, the file size is varied from 10 to 200 Kbytes. For each transmission two decision strategies are supported:

- Always Cheapest network selection (AC)
- Consumer Surplus network selection (CS)

In the case of CS network selection strategy the intelligent node will shortlist the networks based on delivery time prediction to meet T_{c2} . The file is then transmitted over the short-listed RAN which is predicted to maximize the CS.

4.2 Simulation Model

The simulation model was developed using Network Simulator version 2.27 (NS2) [11] with IEEE 802.11b wireless LAN parameter settings (data rate 11 Mbps) and the NOAH (No Ad-hoc routing) extension. The topology used consists of a wired network connected to two Access Points (AP). Each AP has 19 associated wireless nodes. The new application, which simulates a multi-homed terminal with inbuilt CS and AC based network selection strategy, was deployed on two nodes in the overlap area between the RANs (one node attached to each WLAN). The remaining nodes were used to generate background traffic. The wired links have high bandwidth with negligible delay such that end-to-end delay is mainly dependent on the performance of the chosen wireless access network.

After a warm up period of 30 seconds the intelligent wireless node attempts to send the first file. The RAN selection is based on the current user network selection strategy. The file is then sent using FTP through the selected RAN up to the sink server in the wired network. Once the data is transferred the user waits a think-time before requesting transmission of the next file.

4.3 Testing Results

Figure 4 shows the average transfer completion times achieved for all files sent. For users using the CS network selection strategy the average completion time was significantly lower than for those users employing the AC selection strategy. As expected average transfer completion time increases with file size increase, but the rate of increase is steeper for the AC network selection strategy.

Initial results presented in Table II indicate that the proposed CS strategy produces significantly better results for the user over all transferred file sizes. For example a file that is 200 Kbytes long has an average T_c (completion time) over three-fold shorter when using the proposed CS-based strategy than when using the AC strategy.

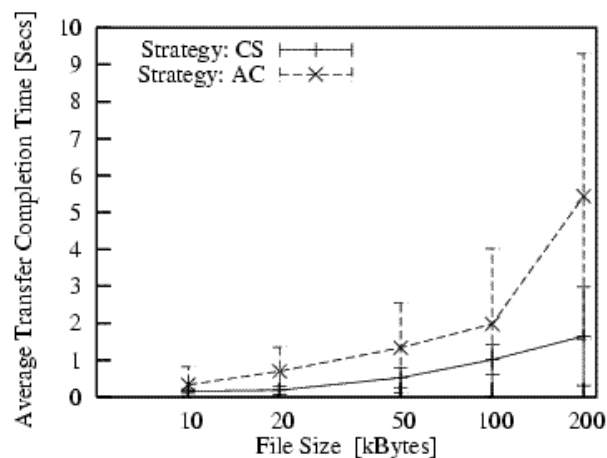


Figure 4. Comparison of Average Transfer Completion Times

TABLE II
AVERAGE T_c FOR AC AND CS NETWORK SELECTION STRATEGIES

File size (Kbps)	Avg. T_c AC (secs)	Avg. T_c CS (secs)	Avg. T_c AC/Avg. T_c CS Ratio
10	0.333	0.155	2.16
20	0.697	0.185	3.77
50	1.334	0.524	2.55
100	1.978	1.013	1.95
200	5.434	1.649	3.29

Most users might be concerned with the average delay. However it is possible that they place a higher importance on the file transfer completing by the deadline. Files which arrive after T_{c2} are worth 0 cent and represent an expense to the user with no return. In Figure 5 the analysis of the number of file transfers that exceed the user frustration threshold T_{c2} is presented. It is found that the proposed CS strategy is highly effective in meeting the delay deadline. On the other hand, in the AC network selection strategy, the user exceeds the completion deadline a lot more. For example, for the 200 Kbyte file size, more than 34% of all transfers do not meet the deadline. The CS strategy on the other hand fails the deadline on less than 1.2% of file transfers for this file size. This improvement is an order of magnitude, and is an example of the real benefit of using the CS-based algorithm when selecting the network.

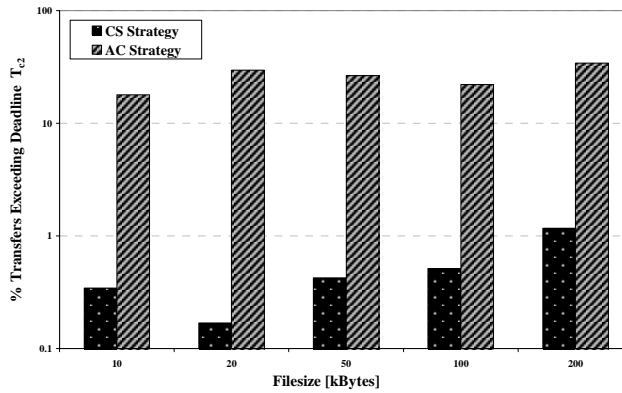


Figure 5. The Percentile Difference in Number of Transfers

5. Future Work

Users in next generation networks will make their network selection decision under the complexity and uncertainty of the dynamic radio environment. They should rely on some criteria to make their choice and in order to evaluate their decision they will look at the consequences of each alternative. For many business decisions the single evaluation measure of money is often used to describe and relate the possible outcomes. However in order to determine the best user utility function along the user's willingness-to-pay their attitude towards taking risks must also be considered. Different individual users will have different preference trends, which may change dynamically with change in context. There are three different categories of risk taking behaviour. The user may be risk adverse, risk neutral or risk seeking, depending on the application and the situation. Kirkwood in [12], discusses many possible shapes for the utility function based on user attitude to risk. Our current user utility function has a shape related to risk neutral user behaviour. Future work will investigate the effect of using utility

functions with either risk seeking or risk adverse related shapes and will compare the performance of the resulting network selection decisions.

6. Conclusions

This paper has presented an overview of the issues facing a user with multiple access networks. It is proposed that the user device will make the decision, as opposed to the networks. This paper concentrates on the decision for the transfer of non real-time data. A selection strategy is proposed based on the benefit that the transfer will give to the user, and this is compared to a more simplistic strategy.

The simulation results presented in this paper show that the proposed strategy (CS) achieves superior results in terms of average transfer completion times compared to the Always Cheapest (AC) strategy. Furthermore, in relation to the time deadline of the file transfer, the CS strategy can be an order of magnitude better than the AC one. It remains to explore how these schemes will deal with other types of utility functions, but it is believed that similar results will be obtained.

REFERENCES

- [1] T. Kanter, "Going Wireless: Enabling an Adaptive and Extensible Environment", *ACM Journal on Mobile Networks and Applications*, Vol. 8, No. 1, pp. 37-50, February 2003.
- [2] J. McNair, F. Zhu, "Vertical Handoffs in Fourth-Generation Multinetwork Environments", *IEEE Wireless Communications*, pp.8-15, June 2004.
- [3] X. Wang, H. Schulzrinne, "An Integrated Resource Negotiation, Pricing, and QoS Adaptation Framework for Multimedia Applications", In *IEEE Journal on Selected Areas in Communications (JSAC)*, vol. 18, December 2000.
- [4] Sajal K. Das, Haitao Lin and Mainak Chatterjee, "An Econometric Model for Resource Management in Competitive Wireless Data Networks", *IEEE Network Magazine*, vol. 18, No. 6, pp. 20-26, November/December 2004.
- [5] J. Ylitalo, T. Jokikyyny, T. Kauppinen, A. J. Tuominen, J. Laine, "Dynamic Network Interface Selection in Multihomed Mobile Hosts", 36th Hawaii International Conference on System Sciences, Hawaii, USA, January 2003.
- [6] G. Le Bodic, J. Irvine, D. Girma, J. Dunlop, "Dynamic 3G Network Selection for Increasing the Competition in the Mobile Communications Market", *IEEE Vehicular Technology Conference*, Boston, USA, September 2000.
- [7] M. Falkner, M. Devetsikiotis, I. Lambadaris, "An Overview of Pricing Concepts for Broadband IP Networks", *IEEE Communications Surveys and Tutorials Q3*, 2000.
- [8] E. Bircher, T. Braun, "An Agent-Based Architecture for Service Discovery and Negotiations in Wireless Networks", 2nd Intl. Conf. on Wired/Wireless Internet Communications, Frankfurt, Germany, February 2004.
- [9] I. Ramani, S. Savage, "SyncScan: Practical Fast Handoff for 802.11 Infrastructure Networks", *Proceedings IEEE Infocom 2005*, Miami, FL, March 2005.
- [10] P. J. Sevcik, "Understanding How Users View Application Performance", *Business Communications Review*, Vol. 32, No. 7, pp. 8-9, 2002.
- [11] C. W. Kirkwood, *Notes on Attitude Toward Risk Taking and the Exponential Utility Function*, 1991 (reissued 1997).
- [12] Network Simulator-2, [Online]. Available: <http://www.isi.edu/nsnam/ns/>