

User- and Application-Centric Multihomed Flow Management

Olivier Mehani^{*†‡§}, Roksana Boreli^{*†}, Michael Maher^{†¶} and Thierry Ernst^{‡§}

^{*}NICTA, Locked Bag 9013, Alexandria, NSW 1435, Australia

Email: {name.surname}@nicta.com.au

[†]University of New South Wales, Sydney, Australia

[‡]CAOR—Centre the Robotique, Mathématiques et Systèmes, Mines ParisTech, France

[§]Imara Project-Team, Inria, 78153, Le Chesnay Cedex, France

Email: {name.surname}@inria.fr

[¶]Reasoning Research Institute, Sydney, Australia

Email: mmaher@cse.unsw.edu.au

Abstract—We consider the problem of network selection and flow distribution for a multihomed mobile device. We argue the benefits of a holistic approach which considers user- and application-centric metrics such as quality, energy consumption and monetary cost, rather than the commonly used network-centric metrics. We thus introduce the *multihomed flow management* problem which combines network selection, flow distribution and application flow awareness. We formulate it as a constrained optimisation problem and compare it to commonly used techniques: single network selection and load balancing. For selected interactive applications, we use empirical network measurements to evaluate the optimal solutions obtained by the three approaches. We show that, by exploiting the flexibility of application parameters, it is possible to achieve the potentially conflicting goals of maintaining high application quality while reducing both the power consumption and cost of network use.

Index Terms—network selection, flow distribution, mobility, multihoming, constrained optimisation problem, quality of experience

I. INTRODUCTION

End-user mobile devices increasingly support multiple interfaces, enabling them to connect to different wireless network technologies. A number of service and network providers may also exist for any specific access technology which, in addition to the advent of User-Provided [25] and Delay-Tolerant [5] Networks, further increase the number of available access network and the need for appropriate selection. The concurrent availability of a number of networks presents both the opportunity and the problem of selecting the most appropriate ones, with the goal of providing “Always Best Connected” [7] devices. A related problem is the distribution of application data flows over the selected networks, for all the applications running on the mobile device.

Common approaches to network selection are based on

the estimated Quality of Service (QoS)¹ parameters on the available networks. The most common criteria include network QoS parameters such as capacity or delay [1], [3], [23], [26], [27], [28], [29], [30]. QoS information can readily be obtained through, *e.g.*, collaborative databases of mobile based measurements [19] or protocols such as the IEEE 802.21 standard [22], [30]. Researchers have also proposed more user-centric criteria which are highly relevant to a mobile user, such as the power consumption [20], [27], [29], or the cost of network use [1], [2], [3], [28]. In some of the aforementioned, the user is also expected to provide policies or preferences [1], [2], [26], [27].

Different approaches are proposed to address the multi-criteria decision problem. [2] uses a combination of fuzzy logic-based decisions and genetic algorithms to derive the weights of each criterion. A heuristic classification algorithm based on the Technique for Order Preference by Similarity to Ideal Solution is used in [3] to compare the available solutions to the ideal expectations. [26] uses the Analytical Hierarchy Process to rank the choices. In [29], the authors approach the network selection as a bin-packing problem. To the best of our knowledge, however, no study has considered constraint programming techniques and the use of dedicated solvers.

The distribution of applications flows over the selected networks is commonly approached as a load-balancing problem. The network QoS characteristics are used as a basis for balancing the total traffic load over the available interfaces [14], [24], [31]. We note that flow distribution can be considered both for the upstream traffic, which can be easily controlled directly by the mobile device and the downstream traffic. For the latter, existing frameworks such as those described in [18], [32] may be used.

We propose a novel holistic approach to solving the problems of network selection and flow distribution, which we call the *multihomed flow management* problem, based

¹In this paper, we use the term Quality of Service loosely to refer to what a network can actually provide in terms of capacity, delays, packet losses, security or other criteria relevant to the applications.

on the criteria of application quality, mobile resource use and price of network service. Our approach directly considers the application quality metrics, rather than indirectly by relying on the network QoS. We argue that QoS-based decisions may not lead to the best user-perceived performance, due to the non-linear relationship between the applications quality and QoS (as exemplified in Fig. 1), and that applications quality predictors should be used instead. A small number of previous works consider application Quality of Experience (QoE) [16] to determine the optimum network selection [21]. However, they use it only as a single global metric, while we recognise quality can vary differently depending on the application, and should therefore be treated with a finer level of granularity. Moreover, to the best of our knowledge, no previous proposal covers both network selection and flow scheduling.

Our proposal includes adaptive variation of the application and protocol parameters. By pre-emptively determining the best set of parameters, we can greatly shorten the adaptation process of the application and underlying transport protocol parameters, among others, to the network conditions. Though this is a cross-layer solution, our proposal is a generic mechanism with a global view of the network stack and will thus avoid the potentially adverse interactions such designs commonly risk [15] when only a limited view is considered.

We model the multihomed flow management problem as a constrained optimisation problem. In this framework, we also model the proposed approach, as well as two commonly used approaches: selection of a single network with highest capacity which is the approach currently used in most smart-phones (*e.g.*, iPhone and Android-based phones) and multihomed load balancing.

We implement the model and approaches in the MiniZinc constraint programming language [17]. This allows us to find the optimal solutions that each approach could yield in a range of scenarios, derived from empirical network measurements. We focus on interactive applications (voice over IP, video conferencing and web browsing) for which the telecommunications community has already defined QoE metrics.

We evaluate the optimal solutions obtained by the three approaches and demonstrate the advantages of our proposal which achieves a significantly better trade-off between the potentially conflicting criteria of applications quality, cost of network use and mobile battery use.

The paper is organised as follows. The constrained optimisation model is presented in Section II. Section III describes the QoE metrics for interactive applications, used in our approach. Section IV presents the evaluation scenarios, with the evaluation of our proposal and comparison with commonly used approaches presented in Section V. We conclude and present future work in Section VI.

TABLE I
SETS AND THEIR OPERATIONS USED TO DEFINE AND MODEL THE MULTIHOMED FLOW MANAGEMENT PROBLEM.

Set of networks N	
None $\in N$	null network to represent unassociated interfaces
Set of interfaces I	
$\vec{A}, \vec{A} = I $	network association vector where $A_i \in N, \forall i \in I$
Set of links $L \subseteq I \times N$	
$QoS(l)$	achievable QoS achievable on link $l \in L$
$Pw(l)$	power consumption of link l
$Pr(l)$	access price of link l
QoS tuple $q = \langle c, r, e, s, \dots \rangle$	
$C(q) = c$	available capacity
$R(q) = r$	round-trip time
e	link error rate
s	security index
\dots	other metrics relevant to an application
Set of flows F	
$\vec{D}, \vec{D} = F $	flow distribution vector where $D_f \in L, \forall f \in F$
$\vec{p}, \vec{p} = F $	application-specific parameters (p_f for flow f)
$Q(f, p_f, q_f)$	quality profile of flow $f \in F$ under QoS q_f
$q_{req}(f, p_f)$	min. required QoS to maximise $Q(f, p_f, q_{req}(f, p_f))$

II. MULTIHOMED FLOW MANAGEMENT

In the remainder of this paper, we collectively refer to application quality, battery use and access price as *overall performance metrics*. We propose to optimise these criteria by deciding on network associations, distribution of flows across links and application parameters.

We model this proposal as a constrained optimisation problem. This section presents the model, the objective function to be optimised, and the structural constraints. The sets and operations are summarised in Table I.

A. The Flow Management Problem

Let I be the set of network interfaces, N the set of all available networks, including the special network None, and $L \subseteq I \times N$ the set of links that can be established between interfaces and networks. At all times, each interface $i \in I$ is associated to a network $n \in N$. This is represented as link $l_i = (i, n)$. Vector \vec{A} of size $|I|$ represents the network associations of all the interfaces $i \in I$; $A_i = n$ when $l_i = (i, n)$.

We define operation $QoS(\cdot) = \langle c, r, e, s, \dots \rangle$ on elements of L which represents the QoS achievable on a given link. Components of $QoS(l)$ include the capacity of the link $C(l) = c$, the round-trip time $R(l) = r$ and the potential error rate e . It can also include indices such as the security level s (*e.g.*, WPA2 would rank better than WEP). We also define two other operations on links, the induced power consumption $Pw(\cdot)$ and the price that some operators charge for use of their network, $Pr(\cdot)$.

Additionally, let F be the set of applications flows which have to be distributed on active interfaces. The performance quality of a flow f can be expressed as $Q(f, p_f, q_f)$ where p_f is a set of application configuration parameters (such as codec or bit-rate) and q_f is the QoS

the flow gets. Examples of functional relations usable as $Q(\cdot)$ are given in Section III for video, voice and web QoE. We denote $q_{\text{req}}(f, p_f)$ the QoS so that $Q(f, p_f, q_{\text{req}}(f, p_f))$ is the highest. This is the requirement for a flow with parameters p_f to perform best. Finally, each flow f must be distributed on one single link, $D_f = l \in L$, where \vec{D} is the flow distribution vector, of size $|F|$.

The flow management problem thus consists in maintaining a high performance quality while keeping low power consumption and access prices. This is achieved by selecting the network association for each interface (possibly turning selected interfaces off), distributing the flows over the active links and adjusting application parameters to the best matching set. This triple objective can be expressed as

$$\max_{\vec{A}, \vec{D}, \vec{p}} \left(\sum_{f \in F} W_f Q(f, p_f, q_{\text{req}}(f, p_f)) - W_b \sum_{i \in I} Pw(l_i) - W_p \sum_{i \in I} Pr(l_i) \right), \quad (1)$$

where the W_* are weighting factors which can be used to scale performance metrics to comparable ranges, and express their relative priority. Also the following structural constraints apply,

$$\left\{ \begin{array}{l} \forall f \in F, \exists i \in I \quad A_i \neq \text{None} \wedge D_f = l_i, \\ \forall i \in I \quad \sum_{f \in F | D_f = l_i} C(q_{\text{req}}(f, p_f)) \leq C(QoS(l_i)). \end{array} \right. \quad (2a) \quad (2b)$$

In other words, (2a) attributes a single active link to each flow and (2b) ensures that the maximal capacity available on each interface is respected.

A noteworthy fact about this model is that, in the process of optimising (1), it derives the QoS that the flow is expected to receive (for example its throughput). This information can be reported as a hint to the transport protocol the flow uses in order to skip its adaptation phases and directly adjust the rate to the selected conditions.

B. Comparison to QoS-based Decisions

To evaluate the potential improvements that could be achieved by the increased awareness of user and application requirements, we compare the proposal to two other mechanisms: 1) selecting only the network with the highest capacity and 2) load balancing flows over all interfaces, with each of the interfaces connected to the highest capacity uplink. The objective functions and additional required constraints for those two mechanisms are described below.

1) *Network Selection*: This first mechanism is identical to what is currently available in consumer devices such as smart-phones (*e.g.*, Android-based or iPhones). These devices will use a Wi-Fi connection in preference to the 3G connection. This default policy is logical as, in general, the Wi-Fi connection is likely to provide a higher capacity at a lower cost of network use.

The network selection problem can thus be represented as

$$\max_{\vec{A}} \sum_{i \in I} C(l_i) \quad (3)$$

$$\text{s.t.} \quad \begin{cases} \exists i \in I \quad A_i \neq \text{None}, \\ \forall j \in I - \{i\} \quad A_j = \text{None}. \end{cases}$$

2) *Load Balancing*: For the load balancing scenario, we consider a device which associates all its interfaces with their respective best networks, chosen by any or other QoS criteria. The current flows are then distributed so that each interface is equally loaded with respect to their available capacity.

To maintain loads roughly equal on all the interfaces of the terminal, we use a formula based on Jain's fairness [13] as an additional optimisation objective. Rather than intrinsic capacity usage, we are interested in load balancing the flow requirements over links according to their capacity. We thus define for each link l a load ratio, $Lr(l) = \sum_{f \in F | D_f = l} C(q_{\text{req}}(p_f)) / C(l)$, which we use in the fairness index,

$$F_r = \frac{(\sum_{i \in I} Lr(l_i))^2}{|I| \sum_{i \in I} Lr(l_i)^2}. \quad (4)$$

This index is 1 when all load ratios are equal, and tends to 0 with increasing unfairness in the load.

The load balancing problem can thus be represented as

$$\max_{\vec{A}, \vec{D}} \left(W_c \sum_{i \in I} C(l_i) + W_f F_r \right), \quad (5)$$

which represents the ideal load-balancing, possibly better than what actual (sub-optimal) algorithms can achieve.

Next, we overview quality metrics defined by the telecommunications community for interactive applications.

III. QOE MODELS BASED ON QOS METRICS

There has been a large body of research and standardisation work in the telecommunications community to provide QoE estimates based on QoS [16]. QoE is defined by the International Telecommunication Union (ITU) as "the overall acceptability of an application or service, as perceived subjectively by the end-user" [11]. The subjective quality is rated by a *Mean Opinion Score* (MOS) value [12]. This section reviews the ITU quality metrics and their non-linear relation to the received QoS.

A. Multimedia Applications

The ITU defines the E-model [9], [10] as the relationship between the multimedia (voice, video) application quality and parameters of the communication link (*e.g.*, equipment, codec, network QoS).

The quality of voice conversations is represented by the value of the *R factor* [9],

$$R = 93.193 - I_s - I_d - I_{e\text{-eff}}, \quad (6)$$

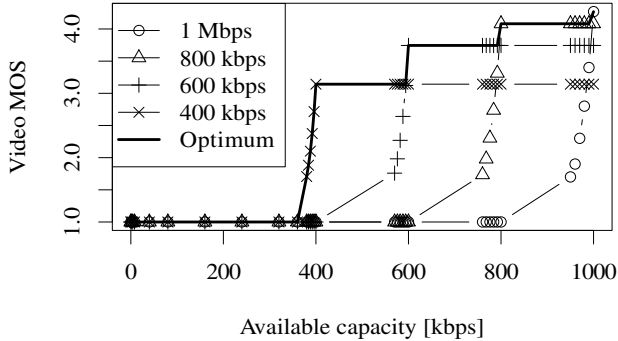


Fig. 1. Quality profiles (V_q) for a video stream encoded with H.264 at various rates for varying network capacities. The optimum MOS is not a linear function of the available capacity. At a chosen bit-rate, a reduction in capacity may make no difference or may make the video unusable.

where I_d is the one-way delay and echo impairment, I_{e-eff} reflects the audio codec and its robustness to losses. I_s is the simultaneous impairment factor, which we ignore hereafter and set to 0 as suggested in [6]. The R factor can be remapped to MOS with a cubic relation [9].

A similar but delay-independent formula is proposed for the video MOS [10],

$$V_q = 1 + I_{coding} \exp\left(\frac{P_{plV}}{D_{PplV}}\right), \quad (7)$$

where I_{coding} is the basic quality of the video codec for a given bit-rate and frame rate, while D_{PplV} is its loss robustness, and P_{plV} is the current loss rate.

Those quality profiles tend to exhibit a non-linear behaviour (see for example Fig. 1 for a video encoded with H.264 at different bit-rates). Without knowledge of the application, reducing its allotted QoS can have adverse consequences on its quality. This strongly supports our hypothesis that application quality metrics are more relevant than raw QoS for flow management.

Application quality information allows more flexibility in the distribution of flows and consideration of possible trade-offs to reduce power consumption and price of access while still providing an acceptable QoE across all applications.

B. Elastic Traffic

ITU recommendation 1030 [8] provides similar metrics for more general IP-based applications. It includes a regression-based estimate of the quality of a web-browsing session depending on the network QoS and the size s of the web page,

$$MOS_{web} = 5 + 4 \cdot \frac{\ln(WeightedST) - \ln(Min)}{\ln(Min) - \ln(Max)}, \quad (8)$$

$$WeightedST = 0.98 \cdot T_3 + 1.76 \cdot T_4, \quad (9)$$

where $T_3 = R(q)$ is the time to display the first elements and $T_4 = s/C(q)$ that to finish all the elements,² respectively, while Min and Max are the minimal expected and maximum accepted times for completion of the request.

Next, we include the presented QoE metrics in the model we introduced in Section II as the $Q(\cdot)$ relations. We then proceed to compare the performance of our proposal (1) to the non application-aware approaches (3) and (5).

IV. EVALUATION SCENARIOS

To evaluate and compare the performance of all three approaches, we implement them in the MiniZinc modelling language [17] and evaluated the various approaches in multiple scenarios. Each scenario differs in the number and type of interfaces, the possible links and their achievable QoS, and the number and type of flows. All parameters come from empirical measurements. The following first details our implementation of the model, then describes the empirical data-sets and explains how they are used to create evaluation scenarios.³

A. MiniZinc Model

The models are relatively straightforward representations of the constraints expressed above. The functional relationships between network and monetary cost; network and interface, and QoS; and network and interface, and power consumption, are all represented as table constraints. Furthermore, the functional relationship between application parameters and QoS and QoE is also expressed as a table constraint. In this latter case the table is a discrete approximation of the function. Consequently, the model has the form of a traditional constraint satisfaction problem (CSP) [4] for which we wish to optimise the objective function.

The MiniZinc language is supported by several different solvers. To run the scenarios we used the default solver. This solver employs the standard constraint programming approach: it uses constructive search, constraint propagation, and branch-and-bound pruning to completely explore the space of possible solutions and find the optimal one.

So far we have not put effort into improving the solving time of the implementation. One possibility is to use more advanced solvers or alternative solving techniques. Another is to refine the way the model is implemented in MiniZinc. Since the thesis of this paper is independent of the method used to optimise the objective function, we leave such improvements for future work.

B. Numerical Parameters

This section describes the numerical parameters used in the implementation of the model or the evaluation scenarios.

²(9) is a simplification of the MOS for an expert user expecting an intermediate session time, discarding the delays due to the initial search request.

³The MiniZinc model as well as the scenarios discussed are available online at <http://www.nicta.com.au/people/mehanio/canso>.

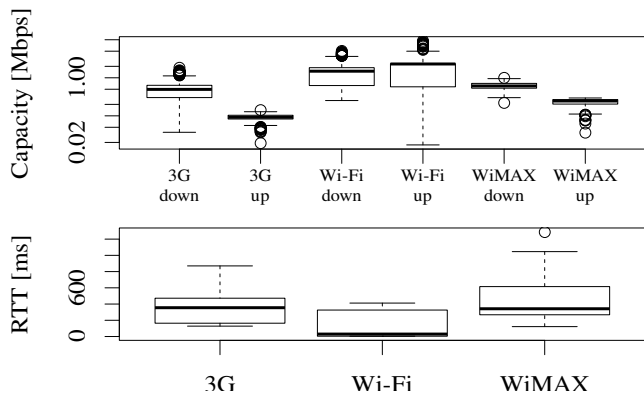


Fig. 2. Links QoS distributions from the parameters data-set used in the synthetic scenarios. Top: upstream and downstream capacity (logarithmic scale), bottom: round-trip time. Boxes show the 25–75% inter-quartile range (*IQR*), the heavier line represents the median of the data set and whiskers extends to the extreme data within $1.5IQR$ from the median. Circles denote outliers.

a) *QoS*: The offered QoS from various free and for-a-fee mobile networks has been measured from several locations in Sydney, Australia, and Bremen, Germany, over a period of a couple of months, to fixed correspondent nodes. Three servers were used as the end points: one on an Australian ADSL2+ line, one on the Aarnet academic network, and one on the French equivalent, Renater. The measurements included down- and upstream capacities (as seen by TCP connections), and round-trip times. Fig. 2 summarises the measured characteristics for each network type.

b) *Battery Consumption*: The battery consumption over 3G and Wi-Fi network interfaces of an HTC phone under different loads and RSSIs has been studied in [20]. One finding is that the power consumption barely varies with the transmission speed. The main drain comes from whether the radio circuit is powered. We re-used the battery measurements from that article for our scenarios and, based on its finding, limited the power usage of links to a fixed ratio of the time the interface is connected to a given network.

As an equivalent data-set for the power consumption of WiMAX adapters was not available, we have used the Wi-Fi measurements as nominal WiMAX values. We believe that this does not qualitatively change the outcome of our comparison. Obtaining WiMAX power usage traces is one of the items for future work.

c) *Access Price*: The pricing terms of the main mobile broadband operators in Australia have been collected in December 2010. There is a wide variety of contract types (timed or quota, plan or prepaid, peak and off-peak periods) and ways to handle excess usage (increased price, data blocks, traffic shaping). Our model does not currently encompass all those different pricing methods. However, as a first approximation, only pricing per connection time has been taken into account. All Wi-Fi networks were considered cost-free.

TABLE II
PARAMETER RANGES USED FOR THE GENERATION OF THE SYNTHETIC SCENARIOS.

Technology	Interfaces		Networks	
	Min	Max	Min	Max
3G	1	1	0	1
Wi-Fi	1	2	0	10
WiMAX	0	1	0	2
Flow type	Min		Max	
VoIP	0		3	
Video	0		3	
Web	5		10	

d) *Quality of experience*: The QoE for video, audio and web flows has been computed for a wide range of supporting QoS. The formulas provided by the ITU for real-time codecs include information about packet loss rate. We do not have direct data about this parameter as our QoS measurements did not include it. Instead, we estimate it based on the available capacity and the flow bit-rates. If $C(q_f) < C(q_{req}(f, p_f))$, a ratio $plr = (C(q_{req}(f, p_f)) - C(q_f)) / C(q_{req}(f, p_f))$ of packets is lost. This ratio is used as the packet loss rate to compute the performance quality of a flow with this reduced capacity as $QoE(f, p_f, (c_f, r_f, 0, \dots)) = QoE(f, p_f, (c_{reqf}, r_f, plr, \dots))$, where $c_f = C(q_f)$, $r_f = R(q_f)$ and $c_{reqf} = C(q_{req}(f, p_f))$. This assumes that packets lost due to too small a capacity are not retransmitted. This is reasonable as the time it takes to retransmit a packet makes it useless for real-time media streaming.

e) *Web Demand*: Unlike real-time constant bit-rate streaming, web traffic does not have a throughput requirement. Rather, as shown in (8), the perceived quality depends on the duration of the transfer, which in turn depends on both the available capacity and the page size. We therefore use the data-set from [20] for the size distribution of web pages.

f) *Priority and Scaling Factors*: Though the optimisation objective functions (1) and (5) include weighting factors W_* for all singular objectives, they are all equal and set to 1 for now. It is the subject of future work to study how to adjust them from user feedback.

C. Generic Scenarios

We first consider a set of generic scenarios which cover a wide range of use cases for the mobile device (such as a smart-phone or vehicular router). Those scenarios are synthetically generated by a random process as follows. The number of interfaces, networks and flows is first determined. Then, the power consumption, cost and QoS of the valid links are randomly chosen from the data-sets. The identifier of the scenario is used as the seed of the pseudo-random number generator in order to allow identical recreation. The parameter ranges for the scenario generation are shown on Table II.

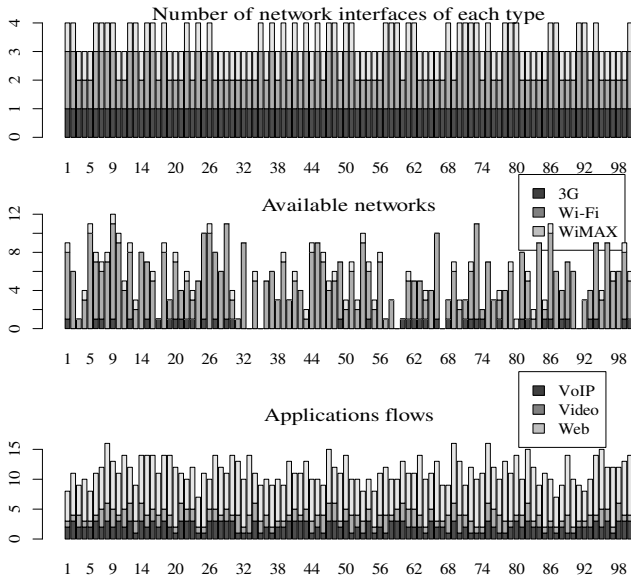


Fig. 3. Scenarios parameters. Top: number of interfaces and available networks of each technology in the studied synthetic scenarios; Bottom: number of flows per type.

Fig. 3 shows the parameters of the scenarios we considered. We arbitrarily chose the first 100, though 5 of them did not have any available networks, reducing the number of significant scenarios to 95. For each scenario, the three approaches have been evaluated by the constraint solver to find the optimal solution each technique could yield. In order to study how the different approaches perform under different demands, the number of flows has been varied from 1 to the total number for each scenario. This allowed us to observe the variation of overall performance metrics, which we discuss in the next section.

D. Typical Smart-phone Use Scenarios

We additionally study the performance of the proposal for more specific scenarios of the smart-phone use case. We consider a two-way video conversation and web browsing. This gives a fixed demand of 2 VoIP flows, 2 video flows, and 3 web sessions. The sizes of the web sessions are taken from the nearest rank from the 40, 50 and 60th percentiles of the data-set. We also limit the scenarios from the previous sub-section to those which have a single Wi-Fi interface, as is currently the case for hand-held mobile devices. The available networks remain the same.

V. RESULTS AND DISCUSSION

We ran the scenarios presented above for our combined approach as well as for the network selection and load balancing schemes. The number of flows from each scenario was varied from 1 to the maximum of each scenario to study the behaviour of the different approaches under various demands. However, the solving models were not implemented with speed of solving in mind. As a result, the time to find the optimal solution increases more than linearly with the number of flows.

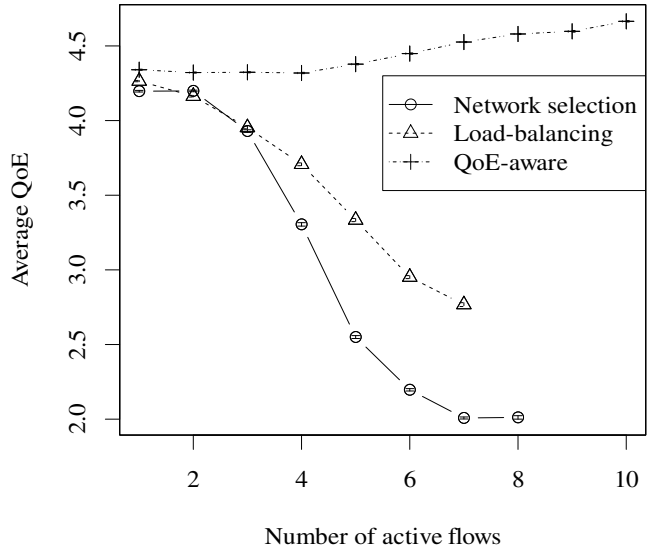


Fig. 4. Average QoE achieved by the three decision schemes. The QoE-aware approach outperforms the two others as it is the only one to manipulate application parameters to match the current QoS.

To avoid unmanageable completion-waiting times, we added a stopping condition in the iteration from 1 to $|F|$ flows if the most recent solution took more than a given time to find.⁴ This stopping criterion has the adverse effect of limiting the number of samples for large numbers of active flows, which increases the confidence interval of the results. However, with fewer than 7 concurrent flows, solving the problem often takes less than 20 s.

For all scenarios, we focus on the average quality over all flows, as well as the power consumption and the price. The latter two metrics have been transposed to units meaningful for a user: consumption of the full battery in %/s, and price in €/s.

A. Generic Scenarios

Fig. 4, 5 and 6 respectively compare the value of the average quality, power consumption and access price depending on the approach. The error bars show the standard error of the results. Only values for which $n_s > 20$ experiments finished within the deadline are plotted.

Fig. 4 shows the variation in the achievable QoE. As could have been expected, the network selection scheme quickly delivers bad quality because it tries to fit all flows over a single link with limited capacity. The load-balancing approach performs better here, provided there is more than one link available for distributing the flows. The QoE-aware decision system, however, manages to maintain the average quality consistently between 4 and 5. It even increases with a larger number of flows, as it becomes

⁴The scenarios were evaluated on a cluster made of 2GHz Xeon machines running Redhat Linux with kernel 2.6.18-92.1.13.el5 #1 SMP. However, the solver did not make use of the multiple cores of the machine, and ran in parallel with other jobs.

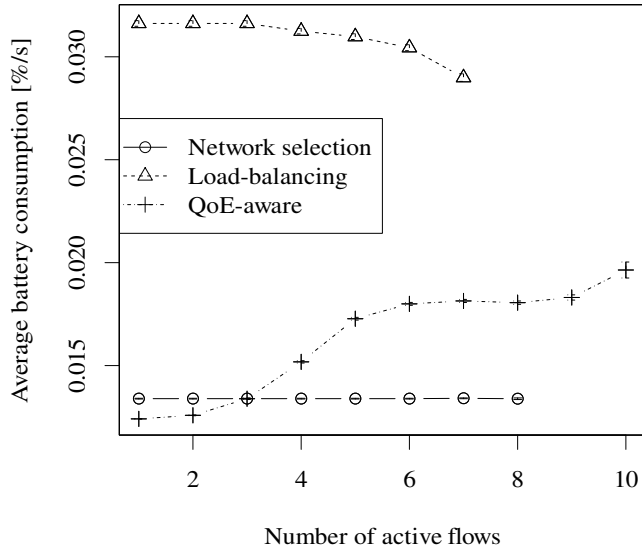


Fig. 5. Battery consumption depending on the number of active flows for each approach. A consumption of 0.03%/s results in a battery lifetime of less than an hour. The apparent decrease in consumption of the load-balancing strategy is biased by the limited number of scenarios with 8 or more flows which could run without hitting the time limit.

effective to enable more interfaces to better support the demand.

The power consumption is shown on Fig. 5. The network selection scheme, using only one interface at a time, usually has the lowest battery consumption. The load balancing, which uses all its interfaces, regardless of the needs, always uses a larger amount of battery. Our proposal has a more dynamic power consumption, which increases as it establishes more links to cater for a higher demand.

Finally, the price is shown on Fig. 6. In the same way as the power consumption, it is directly related to the number of established links. As all the Wi-Fi networks, with the highest capacity, were considered public and free in our scenarios, the network selection scheme unsurprisingly yields a rather low price. The load balancing approach establishes links even on for-a-fee networks, and tries to distribute traffic evenly on them. Using this technique therefore results in higher prices overall. As one of the objectives of our QoE-aware proposal is to keep the price to a minimum, it rarely uses costly network when alternatives exist, even if with lower QoSs, as it can adapt application parameters accordingly.

B. Smart-phone Scenarios

Fig. 7 compares the performance of the proposed performance metrics-aware multihomed flow management technique to the network selection and load balancing approaches with a fixed realistic demand over the subset of scenarios (56) which include a single Wi-Fi interface. We note that the single network selection approach currently implemented in smart-phones provides, on the average,

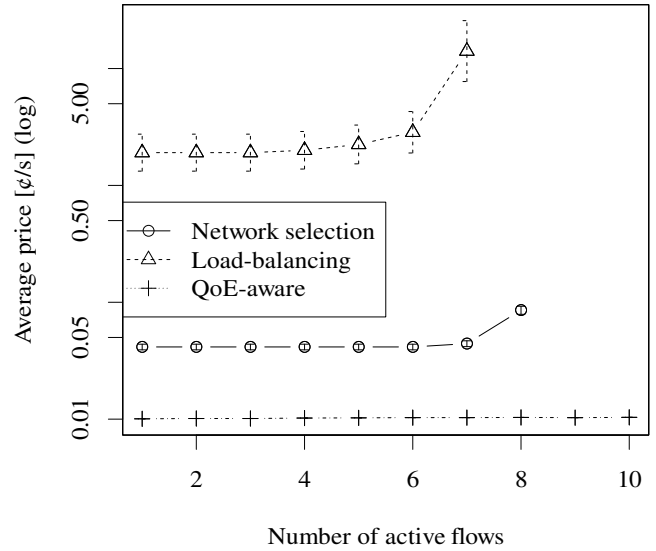


Fig. 6. Price per second of the solutions found by the presented approaches. Only the QoE-aware scheme is mindful of that metric, and tries to keep it as low as possible by using free networks. The apparent increase in price of the load-balancing and network selection strategies is biased by the limited number of scenarios with 8 or more flows which could run without hitting the time limit.

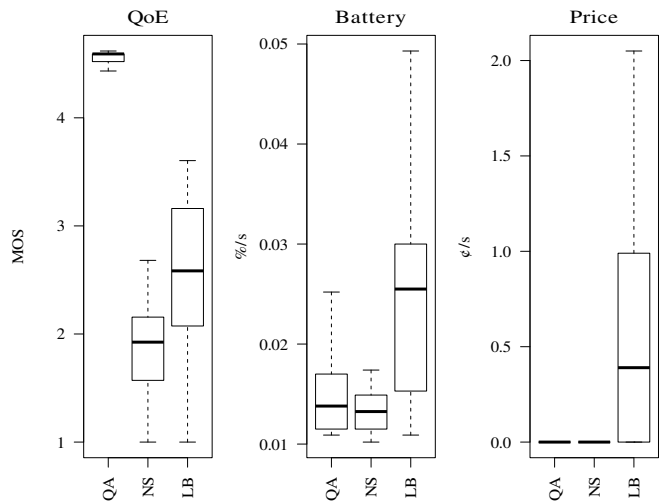


Fig. 7. Distribution of the overall performance metrics yielded by the QoE-aware multihomed flow management (QA), network selection (NS) and load-balancing (LB) approaches in the static demand evaluation. The multihomed flow management approach outperforms the two others in terms of application quality, while keeping reasonably low power consumption and price.

the worst QoE performance and that our proposal, while maintaining a high QoE, manages to keep the average battery consumption in between those of the two others and keep the access price to a minimum.

From those two sets of results, it appears that awareness of application parameters and QoE metrics allows to make

better decisions with respect to which network to connect to, and how to distribute application flows. The possibility to manipulate application parameters based on the knowledge of the QoS they will encounter permits keeping the overall perceived experience high, while maintaining low battery consumption and access costs.

VI. CONCLUSION AND FUTURE WORK

In this paper, we argued that QoS-based network selection is not sufficient to provide a good experience to the user. We introduced a user- and application- aware decision mechanism for multihomed mobile devices. We devised it to select access networks to use for each interface, distribute the application flows over those links and configure the applications to maintain a high quality. In addition, we optimised two other performance metrics, battery consumption and access price.

To evaluate our proposal, we modelled it as a constrained optimisation problem. We compared its performance in terms of overall performance to two common decision mechanisms: the single network selection approach which is currently used in most smart-phones, and multihomed load balancing. Experimental results have shown that the proposed approach out-performs others by supporting high application quality while keeping power consumption and price to a minimum. We also showed how those results remain valid with increasing demand.

It is important to note that the QoE formulas that we used here can be extended with similar relations for other application types. We also believe other similar performance indices can be used for non-interactive applications. This would allow incorporating these into the same decision mechanism.

The results presented in this paper are encouraging and lead to interesting future work. Our aim is to extend the presented proposal further by improving the solving time and incorporating user feedback. We also plan to implement the proposed approach in real devices.

ACKNOWLEDGEMENTS

This research work has been supported by funding from National ICT Australia (NICTA). NICTA is a research organisation funded by Australian Government research initiatives through Australian Research Council (ARC). This work is being carried out in collaboration with the SAIL Project under Grant Agreement Number 257448 of the Seventh Framework Programme of the European Union.

This research was initiated while Michael Maher was employed by NICTA.

The authors would like to warmly thank Henrik Petander for providing his data-sets and Guillaume Jourjon as well as the anonymous reviewers for their insightful comments.

REFERENCES

- [1] E. Adamopoulou, K. Demestichas, A. Koutsorodi, and M. Theologou, "Intelligent access network selection in heterogeneous networks - simulation results," in *ISWCS 2005*, Sep. 2005, pp. 279–283. [Online]. Available: <http://dx.doi.org/10.1109/ISWCS.2005.1547704>
- [2] M. Alkhwilani and A. Ayeshe, "Access network selection based on fuzzy logic and genetic algorithms," *Advances in Artificial Intelligence*, vol. 2008, 2008. [Online]. Available: <http://www.hindawi.com/journals/aai/2008/793058.html>
- [3] F. Bari and V. Leung, "Automated network selection in a heterogeneous wireless network environment," *IEEE Network*, vol. 21, no. 1, pp. 34–40, Jan. 2007. [Online]. Available: <http://dx.doi.org/10.1109/MNET.2007.314536>
- [4] R. Dechter, *Constraint Processing*, May 2003. [Online]. Available: <http://www.worldcat.org/isbn/9781558608900>
- [5] K. Fall, "A delay-tolerant network architecture for challenged internets," in *SIGCOMM 2003*, Aug. 2003, pp. 27–34. [Online]. Available: <http://dx.doi.org/10.1145/863955.863960>
- [6] M. Graubner, P. S. Mogre, R. Steinmetz, and T. Lorenzen, "A new QoE model and evaluation method for broadcast audio contribution over IP," in *NOSSDAV 2010*, Jun. 2010, pp. 57–62. [Online]. Available: <http://dx.doi.org/10.1145/1806565.1806581>
- [7] E. Gustafsson and A. Jonsson, "Always best connected," *IEEE Wireless Communications*, vol. 10, no. 1, pp. 49–55, Feb. 2003. [Online]. Available: <http://dx.doi.org/10.1109/MWC.2003.1182111>
- [8] ITU-T Recommendation G.1030, "Estimating end-to-end performance in IP networks for data applications," ITU-T SG12, May 2006. [Online]. Available: <http://www.itu.int/rec/T-REC-G.1030-200511-1/en>
- [9] ITU-T Recommendation G.107, "The E-model, a computational model for use in transmission planning," ITU-T SG12, Mar. 2005. [Online]. Available: <http://www.itu.int/rec/T-REC-G.107-200904-I/en>
- [10] ITU-T Recommendation G.1070, "Opinion model for video-telephony applications," ITU-T SG12, Apr. 2007. [Online]. Available: <http://www.itu.int/rec/T-REC-G.1070-200704-I/en>
- [11] ITU-T Recommendation P.10/G.100 Amendment 2, "New definitions for inclusion in recommendation P.10/G.100," ITU-T SG12, Jul. 2008. [Online]. Available: <http://www.itu.int/rec/T-REC-P.10-200807-1!Amd2/en>
- [12] ITU-T Recommendation P.800, "Methods for subjective determination of transmission quality," ITU-T SG12, Aug. 1996. [Online]. Available: <http://www.itu.int/rec/T-REC-P.800-199608-I/en>
- [13] R. K. Jain, D.-M. W. Chiu, and W. R. Hawe, "A quantitative measure of fairness and discrimination for resource allocation in shared computer systems," Tech. Rep. DEC-TR-301, Sep. 1984. [Online]. Available: <http://www1.cse.wustl.edu/~jain/papers/ftp/fairness.pdf>
- [14] S. Kandula, K. C. Lin, T. Badirkhanli, and D. Katabi, "FatVAP: Aggregating AP backhaul capacity to maximize throughput," in *NSDI 2008*, Apr. 2008. [Online]. Available: http://www.usenix.org/events/nsdi08/tech/full_papers/kandula/kandula.pdf
- [15] V. Kawadia and P. R. Kumar, "A cautionary perspective on cross-layer design," *IEEE Wireless Communications*, vol. 12, no. 1, pp. 3–11, Feb. 2005. [Online]. Available: <http://dx.doi.org/10.1109/MWC.2005.1404568>
- [16] K. Kilkki, "Quality of experience in communications ecosystem," *Journal of Universal Computer Science*, vol. 14, no. 5, pp. 615–624, Mar. 2008. [Online]. Available: http://www.jucs.org/jucs_14_5/quality_of_experience_in/jucs_14_05_0615_0624_kilkki.pdf
- [17] K. Marriott, N. Nethercote, R. Rafeh, P. J. Stuckey, M. Garcia de la Banda, and M. Wallace, "The design of the Zinc modelling language," *Constraints*, vol. 13, no. 3, pp. 229–267, Sep. 2008. [Online]. Available: <http://www.cs.mu.oz.au/~pjs/papers/constraints08b.pdf>
- [18] K. Mitsuya, R. Kuntz, S. Sugimoto, R. Wakikawa, and J. Murai, "A policy management framework for flow distribution on multihomed end nodes," in *MobiArch 2007*,

- Aug. 2007. [Online]. Available: <http://www.sfc.wide.ad.jp/~mitsuya/PAPERS/mitsuya-mobiarch07.pdf>
- [19] J. Pang, B. Greenstein, M. Kaminsky, D. McCoy, and S. Seshan, "Wifi-Reports: Improving wireless network selection with collaboration," in *MobiSys 2009*, Jun. 2009, pp. 123–136. [Online]. Available: <http://www.cs.cmu.edu/~jeffpang/papers/mobisys09-wifireports.pdf>
- [20] H. Petander, "Energy-aware network selection using traffic estimation," in *MICNET 2009*, Sep. 2009, pp. 55–60. [Online]. Available: http://www.nicta.com.au/research/research_publications/show?id=2159
- [21] K. Piamrat, C. Viho, A. Ksentini, and J.-M. Bonnin, "QoE-aware network selection in wireless heterogeneous networks," Tech. Rep. RR-7282, May 2010. [Online]. Available: http://hal.inria.fr/docs/00/48/26/81/PDF/RR_7282.pdf
- [22] E. Piri and K. Pentikousis, "IEEE 802.21," *The Internet Protocol Journal*, vol. 12, no. 2, pp. 7–27, Jun. 2009. [Online]. Available: http://www.cisco.com/web/about/ac123/ac147/archived_issues/ipj_12-2/122_ieee.html
- [23] I. Psaras and L. Mamas, "On demand connectivity sharing: Queuing management and load balancing for user-provided networks," *Computer Networks*, vol. 55, no. 2, pp. 399–414, Feb. 2011. [Online]. Available: <http://www.ee.ucl.ac.uk/~uceeips/comnet-upns-ipsaras.pdf>
- [24] J. Singh, T. Alpcan, P. Agrawal, and V. Sharma, "A Markov decision process based flow assignment framework for heterogeneous network access," *Wireless Networks*, vol. 16, no. 2, pp. 481–495, Feb. 2010. [Online]. Available: <http://www.tansu.alpcan.org/papers/WinetRevised2.pdf>
- [25] R. Sofia and P. Mendes, "User-provided networks: Consumer as provider," *IEEE Communications Magazine*, vol. 46, no. 12, pp. 86–91, Dec. 2008. [Online]. Available: <http://dx.doi.org/10.1109/MCOM.2008.4689212>
- [26] Q. Song and A. Jamalipour, "Network selection in an integrated wireless LAN and UMTS environment using mathematical modeling and computing techniques," *IEEE Wireless Communications*, vol. 12, no. 3, pp. 42–48, Jun. 2005. [Online]. Available: <http://dx.doi.org/10.1109/MWC.2005.1452853>
- [27] H. J. Wang, R. H. Katz, and J. Giese, "Policy-enabled handoffs across heterogeneous wireless networks," in *WMCSA 1999*, Feb. 1999, pp. 51–60. [Online]. Available: <http://dx.doi.org/10.1109/MCSA.1999.749277>
- [28] A. L. Wilson, A. Lenaghan, and R. Malyan, "Optimising wireless access network selection to maintain QoS in heterogeneous wireless environments," in *WPMC 2005*, Sep. 2005, pp. 1236–1240. [Online]. Available: http://www.wpmc-archives.com/site/user_login.php?&menu_state=inp&essay_id=700
- [29] B. Xing and N. Venkatasubramanian, "Multi-constraint dynamic access selection in always best connected networks," in *MobiQuitous 2005*, Jul. 2005, pp. 56–64. [Online]. Available: <http://dx.doi.org/10.1109/MOBIQUITOUS.2005.39>
- [30] T. A. Yahiya and H. Chaouchi, "An optimized handover decision for heterogenous wireless networks," in *PM2HW2N 2009*, Oct. 2009, pp. 137–142. [Online]. Available: <http://dx.doi.org/10.1145/1641913.1641933>
- [31] J. Yao, S. S. Kanhere, and M. Hassan, "Geo-intelligent traffic scheduling for multi-homed on-board networks," in *MobiArch 2009*, Jun. 2009. [Online]. Available: <http://www.cse.unsw.edu.au/~salilk/papers/conferences/MOBIARCH2009.pdf>
- [32] J. Ylitalo, T. Jokikyyny, T. Kauppinen, A. J. Tuominen, and J. Laine, "Dynamic network interface selection in multihomed mobile hosts," in *HICSS-36*, Jan. 2003. [Online]. Available: <http://dx.doi.org/10.1109/HICSS.2003.1174876>