A Utility-based Resource and Network Assignment Framework for Heterogeneous Mobile Networks

Ilias Tsompanidis*‡, Ahmed H. Zahran*†, Cormac J. Sreenan*‡
*Department of Computer Science, University College Cork, Ireland
†Department of Electronics and Electrical Communications, Cairo University, Egypt
‡The Telecommunications Research Centre (CTVR), Ireland
Email: {it1,a.zahran,cjs}@cs.ucc.ie

***Abstract*—Network utility theory has been extensively employed for resource management purposes. However, traditional utility functions cannot support attributes that affect the resource allocation, such as mobility or more advanced traffic models. Especially in the context of a heterogeneous wireless network, transient parameters can have varying effects on each access network type. This work proposes a new utility function that can support multiple design requirements for mobile networks including advanced traffic models, user classes, handover and session priorities. We integrate the new utility function with the Super Base Station framework and devise a novel trigger-based network and resource assignment framework that efficiently copes with the complexity of a heterogeneous wireless network. Our simulation results show that the proposed sub-optimal trigger-based framework performs equally well as the complex optimal scheme.**

I. Introduction

The prevalence of multi-interface mobile devices has become more apparent in the last years. Modern smartphones and tablets utilise multiple radio technologies in order to keep the users connected to the Internet, as different types of network infrastructure are available, within and across countries. It is now a common practice for mobile operators to support the rollout of their 4G LTE network with the pre-existing 3G and 2G coverage. Some operators may also complement the service with WiFi hotspots, synthesising a highly heterogeneous wireless network.

In such heterogeneous setting, mobile devices typically choose the network access technology based on a predefined handover policy [1]. Additionally, the Points of Attachment (PoAs) individually manage their resources, in a distributed fashion, without taking advantage of the coverage overlap between them. Hence, the utility of both network operators and users is not optimised. To achieve this goal, a joint centralised framework is required to improve the system resource utilisation and user perceived quality of service (QoS) [2].

Several centralised schemes are proposed for network selection and resource allocation. In [3], the authors propose a centralised, utility-based resource allocation scheme for mixed traffic. A centralised network selection scheme is presented to optimise the users' QoS in [4]. In [5] a joint resource allocation and network selection scheme is proposed for static users. In [3], [4] user utility is employed to achieve their goals. However, their utility models are based on standard functions; e.g., logarithmic and sigmoid functions, or discrete thresholds.

In this paper, we propose a centralised framework for joint, utility-based resource and network assignment (JURNA) in heterogeneous mobile networks. In our framework, we devise a novel utility function to accommodate different design goals including supporting different user classes, avoid unnecessary handovers, and prioritising active users over new connections. Additionally, we introduce the new utility function to Super BS framework to enable its applicability to mobile systems. Furthermore, we propose a trigger-based strategy to ensure the scalability of the joint problem solution. Our performance evaluation shows that our developed framework achieves per-class throughput and fairness similar to those achieved by the optimal solution. We also show that our framework achieves such performance at a significantly reduced handover and relatively improved resource utilisation rate.

The remainder of this paper is organised as follows. We present the related work in Section II. Section III discusses the simulated network model. Section IV introduces JURNA framework components including the design of a new utility function and trigger-based assignment problem. Our performance evaluation is presented in Section V followed by our conclusion in Section VI.

II. Related Work

Following Kelly's seminal paper on Network Utility Maximisation (NUM) [6], extensive research work targeted numerous aspects of resource sharing and management. An up-to-date survey of methodologies applied on network selection on Heterogeneous Wireless Networks (HWN) is presented by Wang and Kuo [7], identifying utility theory as the most popular. Utility refers to the satisfaction the user derives from obtaining the service or goods in question. Most research work based on NUM considers that best-effort traffic is modeled by monotonously increasing concave functions, and QoS-bound traffic by convex-concave functions, e.g., logarithmic and sigmoidal utility functions ([3], [6], [8], [9]).

In wireless networks, optimal resource allocation in a single Point of Access (PoA) is now approached by heuristic algorithms that can provide realtime solutions. Abdelhadi, et al., [10] devised a pair of centralised and distributed algorithms that handle both best-effort and real-time traffic in a proportionally fair way. On the other hand, Chen, et al., [3] model best-effort traffic also with a sigmoidal utility, with the slope and inflection point set accordingly, towards a *unified utility function*. Regardless of the approach, the amount of research in this field allows us to assume that it is feasible to achieve near-optimal utility fair resource allocation in a single PoA.

Nguyen-Vuong, et al., focused on utility models in the context of access network selection [8]. They employed a sigmoidal function for different attributes of the access networks (e.g., cost, QoS, and network load), and multiplied them to form a multi-criteria utility. Srivastava and Bullo [11] studied a family of knapsack problems with sigmoid utility that implement a linear cost factor, named *penalty*. They provide feasible solution algorithms, albeit their approach is strictly mathematical and is not applied on practical problems.

Utility theory is also applied in HWN by combining throughput and cost utilities for network selection and resource allocation (e.g., [12]). A NUM problem is solved for allocating bandwidth to the service types present in the network, which is then shared between the users according to bankruptcy game theory. This approach assumes that different access networks impose different monetary costs on the user, whereas we assume the real-world paradigm of flat-rate user subscriptions in single-operator HWNs and introduce a pure utility-based joint network and resource assignment framework.

III. Network Model

We consider a HWN consisting of PoAs with partially overlapping coverage. The bandwidth allocated to every PoA s, denoted as C_s, is considered known and constant. Each PoA offers a number of nominal connection data rates, corresponding to adaptive modulation and coding (AMC) techniques used in different technologies. For each client i, the current maximum achievable throughput $\widehat{t_{i,s}}$ via PoA s is assumed to be known. Each client has minimum and maximum throughput demands, denoted as D_i^{min} and D_i^{max}, respectively.

When a client connects to a PoA, a portion of its available bandwidth is allocated to that client. It is worth noting that this allocated bandwidth, denoted as $b_{i,s}$, contributes to the user throughput t_i, according to $R_{i,s}$, the spectral efficiency of a user at each PoA. Typically, the spectral efficiency depends on the client-PoA channel condition and the adopted AMC.

IV. JURNA Framework

In this section, we first present our JURNA problem formulation followed by important design aspects including the design of the utility function and the reduction of the complexity of its operation.

A. Problem Statement

Our objective is to maximize the sum of users' utility subject to constraints on the available resources and user demand. This program can be expressed as:

$$\underset{t_i}{\text{Maximise}} \sum_i \sum_s a_{i,s} U(s, t_i) \tag{1a}$$

subject to:

$$t_i = \sum_s a_{i,s} R_{i,s} b_{i,s} \tag{1b}$$

$$\sum_s a_{i,s} \le 1, \forall i \tag{1c}$$

$$\sum_i b_{i,s} a_{i,s} \le C_s, \forall s \tag{1d}$$

$$b_{i,s} \ge 0, \forall s, i \tag{1e}$$

$$a_{i,s} R_{i,s} b_{i,s} \le \widehat{t_{i,s}}, \forall s, i \tag{1f}$$

$$t_i \le D_i^{max}, \forall i \tag{1g}$$

$$a_{i,s} = \begin{cases} 1, \text{ if } b_{i,s} > 0 \\ 0, \text{ otherwise} \end{cases} \tag{1h}$$

The solution of this program provides the assignment variables $a_{i,s}$ that defines user-PoA associations and $b_{i,s}$ the allocated resources per user at their respective PoA. A client receiving less throughput that his demand should not lead to an infeasible solution, thus, D^{min} is considered in the utility function and not as an optimisation constraint. We now present the new utility function used in Eq. 1a.

B. Utility Function Design

Concave throughput-utility functions are widely used to model the appreciation of a basic level of throughput to be quantitatively more than an equal increase thereof. This property entails that the maximum total utility is achieved with a fair allocation of throughput among a number of similar clients. Supporting multiple client classes is possible with different utility curves for each class. We define the utility of client i in a priority class as [13]:

$$f_{P_i}(t_i) = P_i \ln\left(\frac{e-1}{P_i} t_i + 1\right), \tag{2}$$

where P_i is a class-specific tuning parameter of the utility function curvature. Note that the larger P_i the higher the user priority. Additionally, this utility definition ensures fairness between clients of the same class.

In heterogeneous mobile networks, there exist several design requirements in addition to fair resource assignment and class differentiation, including the prioritisation of active connections over new ones and avoidance of unnecessary handovers which may represent a huge burden. We implement these attributes by introducing *handicaps* to the utility, i.e., selectively changing the utility for any PoA-client pair as presented below. Hence, we define a PoA-client utility function, denoted as $U(s, t)$, as:

$$U(s, t) = u(t) - h_{HA}(s, t) - h_{MA}(s, t) - h_{AE}(s, t), \tag{3}$$

where $u(t)$ is the base utility, $h_{HA}(s, t)$ is the handicap for avoiding unnecessary handovers, $h_{MA}(s, t)$ is the handicap for deprioritising short-range PoA under mobility, and $h_{AE}(s, t)$ is the handicap for considering the Activity Endowment. It is worth noting that this utility function is a pairwise relation defined between individual clients and PoAs, but the user index is dropped for presentation simplicity. The design of every component in Eq. 3 is presented below.

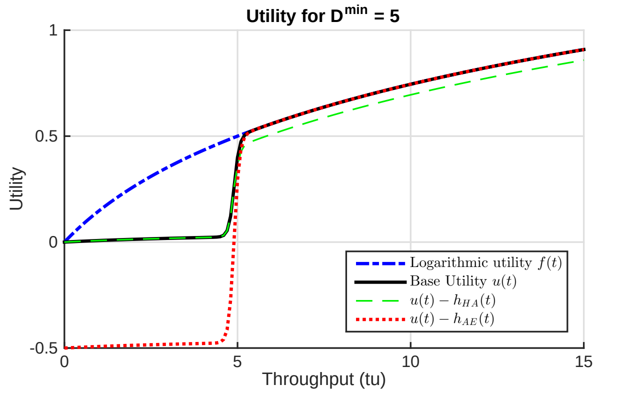

Fig. 1. The base utility function as proposed and the effects of the Endowment Effect and Handover Avoidance handicaps.

1) Base utility: In this work, the utility function is expressed as a product of two components and is expressed as:

$$u(t) = S(t)f(t), \tag{4}$$

where $f(t)$ is a typical concave utility function that also includes user priority information as in Eq. 2 and $S(t)$ is introduced in order to support D^{min} organically in the utility function, such that the utility for throughput $t < D^{min}$ should tend to 0. To attain such mode, we use a sigmoidal logistic function expressed as:

$$S(t) = \frac{A}{(1 + e^{-\omega(t - t^{sp})})} + 1 - A \tag{5}$$

The parameters of Eq. 5 are selected so that $u(t)$ complies with the following set of rules:

1) $u(t)$ has to be strictly increasing.
2) $u(D^{min}) \geq 0.95 f(D^{min})$, so that for $t > D^{min}$, $u(t) \approx f(t)$.
3) The slope angle has to be sharp and proportional to D^{min}, so that $u(0.95D^{min}) \leq 0.05u(D^{min})$.

Hence, the slope point is set to $t^{sp} = 0.98D^{min}$ and the angle to a steep $\omega = 65/t^{sp}$. For D^{min} support amplitude is set to $A = 0.95$ and its lower asymptote is set to $1 - A$, so that $S(-\infty) = 0.05$ and $S(+\infty) = 1$. Figure 1 illustrates the impact of $S(t)$ on a typical logarithmic utility assuming a minimum throughput of 5 throughput units (tu).

This novel approach combines three different attributes into a single utility function, including the concave utility, supporting user classes, and the minimum demand.

2) Handover Avoidance (HA): In heterogeneous mobile networks, users perform frequent vertical and horizontal handovers. The base utility function is agnostic to the overhead associated with handovers and may lead to additional handovers to optimise the resource utilisation. Many of these latter handovers may be performed even when the user is static and some may not result in improving the utility of the user. To alleviate this issue, we introduce a handicap to prioritise the current PoA, by slightly reducing the utility of other PoAs. Hence, two different PoA that can offer the same throughput translate to different utilities, depending on the current connection of the user. Note that this handicap converts the NUM program from a stateless to a stateful problem in which the state is implied in the definition of client-PoA utility.

The HA handicap is defined to comply with the following rule: handover to a new PoA if it can offer at least φ times better throughput than the current one. We further utilise this facility by considering different handicaps for horizontal and vertical handovers, under the assumption that the former are more preferable in terms of handover time and complexity. A multiplier δ is used for this purpose, with indicative values of $\delta_s = 1$ and $\delta_s = 2$, for a PoA s that requires a horizontal and a vertical handover, respectively. For the current PoA, $\delta_s = 0$. For a current throughput t^{cur}, h_{HA} is calculated as:

$$h_{HA}(s, t) \overset{A=1}{=} S(t)\delta_s \left(u(\varphi t^{cur}) - u(t^{cur})\right) \tag{6}$$

The multiplication with $S(t)$ ensures that $h_{HA}(x, 0) = 0$, as it makes no difference to the user whether a handover occurred, if he is blocked. The combined base utility with HA handicap is illustrated in Figure 1.

3) Mobility Assistance (MA): For moving clients, the expected duration of connectivity with a PoA depends on the user speed and the range of the PoA. It is a common design goal to discourage handovers to short-range PoAs while the client device is moving at relatively high speed, as such connections are more likely to disconnect in the immediate future.

Hence, we introduce Mobility Assistance (MA) handicap, a negative utility to deprioritise short-range PoA for mobile clients, estimated as:

$$h_{MA}(s, t) \overset{A=1}{=} S(t)\mu_s \left(u(\varphi t_c) - u(t_c)\right), \tag{7}$$

where μ_s is a factor reversely proportional to the range r_s of PoA s, normalised against the maximum range of all PoA types in the HWN, expressed as:

$$\mu_s = 1 - \frac{r_s}{\max\limits_{s}(r_s)} \tag{8}$$

4) Activity Endowment (AE): It is widely accepted that new sessions should not be accommodated at the expense of ongoing sessions. Hence, we introduce an AE of active user utility to increase its priority over new users[1]. AE can be considered as another handicap h_{AE} of active users to the system and is expressed as:

$$h_{AE}(s, t) = \gamma(1 - S(t, D^{min}))u(D^{min}), \tag{9}$$

where γ can be used to finetune the additional endowment utility. Figure 1 illustrates the impact of changes in the user utility function by introducing AE. Note that h_{AE}, as defined in Eq. 9 supports negative utility values by design when an active client is forced to terminate, while maintaining the same utility of the active user when its assigned resources are kept unchanged. Hence, $U(t)$ is not lower-bound at 0. Since $u(t)$ is logarithmic-based, $U(t)$ also has no upper bound.

[1]The presented AE concept is inspired by the endowment effect in Economics. This principle indicates that a person requires a compensation greater than what he paid in order to abandon the ownership of a good.

C. Triggered JURNA

Optimising the resource allocation in heterogeneous mobile networks mandates frequent changes in resource and network assignment in response to variations in session, channel and mobility dynamics. Achieving this goal in realtime is impractical due to the immense problem complexity. Hence, we propose *triggered JURNA* as a practical approach for resource management.

In triggered JURNA, resource and user reassignment is optimized periodically instead of per every network change. Within this period, the network and resource assignment are performed based on a predefined set of actions. Hence, the system would be operating in a sub-optimal mode and the periodic evaluation enables avoiding long-standing non optimal allocations caused by the defined actions. In the following, we present the main triggers and the associated actions assuming the IRM resource allocation heuristic introduced in [5]. In IRM, an approximation of the client resource allocation is estimated first. Then, this estimation is used to drive network selection and the actual resource allocation.

D. JURNA Triggers

First we start with triggers associated with session dynamics, i.e., session arrival and departure. For departing sessions, we delay resource reshuffling until the periodic run of the resource allocation scheme. We assume a possible sub-optimal system operation as the freed resources would be allocated to other users if needed. Similarly, the system operates in a sub-optimal mode when new sessions arrive. This is attained by pre-associating inactive users with a PoA at the previous periodic run instant. During that run, we consider a class-dependent average demand $\overline{D^{avg}}$, for inactive clients. The client is assigned the BS that is able to better facilitate its $\overline{D^{avg}}$ when it becomes active. Note that no resources are reserved for inactive clients, and that this scheme is supported by a special feature in IRM framework; specifically, supporting negative residual capacity at PoAs. This feature allows the evaluation of resource allocations beyond the actual node capacity.

The second trigger class includes mobility-related events. These events trigger actions that are designed to operate the system sub-optimally until the next periodic run for global resource allocation. Some of these actions are associated with a *scope* limited to the involved clients or PoAs. Thus, the scope selectively enables handovers for particular clients, while resource allocation is managed in a per-PoA basis.

Depending on how wide the allocation reassessment should be, two scopes are defined: *Client* and *PoA*. The Client scope only allows handovers for the client involved in the trigger. The PoA scope allows handovers for all attached clients of the PoA involved in the trigger.

We now define the triggers, describing the events causing them and the subsequent actions the system takes.

1) **Trigger:** Client-Mobile.
   Event: Client movement is identified.
   Action: Flag client as mobile, and enable h_{MA}.
   Scope: N/A.
2) **Trigger:** Client-Stationary.
   Event: The Mobility timer expires.
   Action: Flag client as stationary, and disable h_{MA}.
   Scope: N/A.
3) **Trigger:** Connection-Going-Down.
   Event: Link condition deterioration, e.g., client leaving the active PoA.
   Action: Handover this client.
   Scope: Client.
4) **Trigger:** Bad-Connection.
   Event: Client is under-performing due to overestimated connection quality.
   Action: Re-evaluate link data rate, adjust client handicaps, reshuffle client / resources.
   Scope: Client.
5) **Trigger:** PoA-overload.
   Event: PoA utilised bandwidth exceeds 90%, and is higher than the RMA allocation.
   Action: Reshuffle clients.
   Scope: PoA.

PoA-overload is expected to either result in clients handed over to other PoAs, or identifying the high utilisation to be normal under the current network load. It should be noted that there are more events that could warranty some network side response, however, they are covered with the scheme introduced before. For example, when a client moves into the coverage area of a PoA, the network does not consider a handover at that point. If necessary, the client will be handed over by a PoA-overload, Connection-Going-Down, or Bad-Connection trigger, or during the periodic run. Alternatively, it would be handed over to another PoA during the periodic run if this move is expected to improve the total system utility.

V. Performance Analysis

A. Simulation Setup

We simulate a network occupying a square area of 2 km^2, as shown in Figure 2. The network density is derived from the Nokia dataset of Lausanne [14], [15], with 1.75/km^2 for macro-cell Base Stations (BS), and 20/km^2 for WiFi Access Points (AP). A honeycomb topology is used for 4 macro-cells, for the co-located 4G and 3G BS. The BS range is set to 600 m, to allow partial overlapping. Furthermore, 40 AP were placed uniformly on the map, and their range is 100 m.

The clients are placed randomly on the map, and their movement is dictated by the Truncated Levy Walk model [16]. Their mobility trace can be seen in Figure 2. The position, movement, traffic patterns and demands of the clients are different for each simulation run.

Clients are equally split into three priority classes: Bronze, Silver, and Gold. The traffic demand of the users is assumed to be uniformly selected from $[1, 15]$, $[3, 20]$, and $[6, 30]$ tu for Bronze, Silver, and Gold classes, respectively. The creation of data traffic sessions is described by an ON-OFF model [17]. The ON time follows a Weibull distribution (80, 0.4), with $\overline{d_{ON}} = 265$s. The OFF time is modeled with a

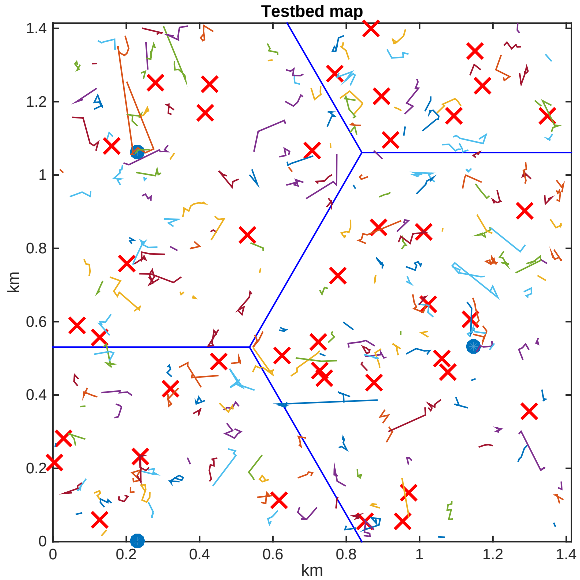

Fig. 2. BS positions in a 2 km^2 map. 3G and 4G are co-located in the centre of each cell, while WiFi APs are placed uniformly (x). The mobility trajectories of the clients are also shown.

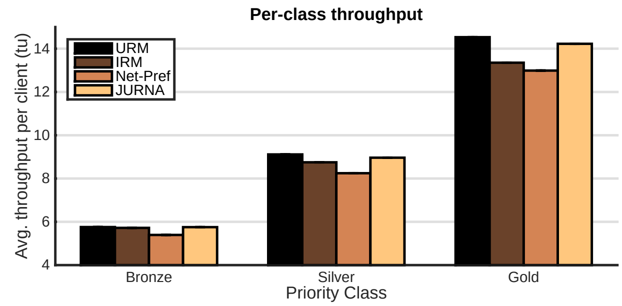

Fig. 3. Average throughput.

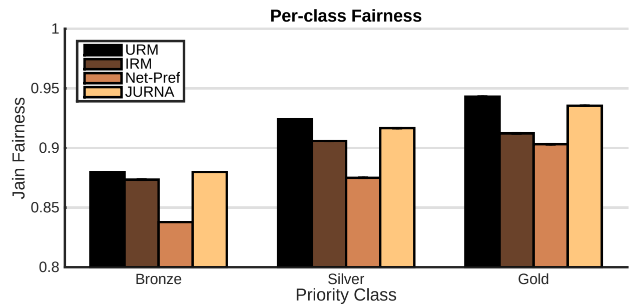

Fig. 4. Fairness (with SuperBS as optimal).

Pareto distribution ($K = 1.3073$, $\sigma = 11.48$, $\theta = 0$), with $\overline{d_{OFF}} = 72$s. The number of clients is selected in accordance with their traffic requirements and the total network capacity. More specifically, the number of clients is 261, so that the average traffic can fully utilise the network.

We evaluate the performance of JURNA with a 2-minute IRM period and compare it with the Network-preference (Net-Pref) strategy, similar to how mobile devices operate today. Under Net-Pref, each client prioritises the network types (WiFi $>$ 4G $>$ 3G), and performs a vertical handover to a better type when available. Horizontal handovers or vertical to worse PoA types are performed when connectivity with the currently associated PoA is lost. When choosing a PoA, the client associates with the one of the best available network type with the highest RSS. For consistency, all approaches assume that PoAs internally use URM to manage their resources.

These two approaches are compared against the utility-optimal URM [13] and the heuristic-driven IRM [5], albeit with the utility function $U(t)$ as defined in Eq. 3. We also employ SuperBS [5] to compute the throughput that each client would get in a fictitious single PoA equivalent to the HWN, as a benchmark for fairness. The proposed algorithms are evaluated with MATLAB simulations and fmincon is used to solve mixed non-linear programs. We consider several performance metrics including average throughput, and blocking probability. A client is considered to be blocked if he is active and his throughput is less than D^{min}. The shown results are based on the average of 100 two-minute simulation runs.

B. Performance Results

Figure 3 shows that JURNA closely achieves a per-class throughput attained by the optimal URM. It also shows that JURNA achieves a per-class throughput gain of 8-12% in comparison to the typical Net-Pref scheme. Similarly, Figure 4 shows that JURNA attains a similar level of fairness to URM. The figure also suggests the superiority of JURNA fairness in comparison to Net-Pref scheme. Further, Figure 5 indicates that JURNA efficiently utilises the available resources by attaining a similar utilisation, in comparison to URM, for both 4G and WiFi networks and a slight reduction in 3G network usage. The figure also suggests that Net-Pref, under-utilises the 3G PoAs, as expected. In fact it does not use them at all, since they are overlapped by 4G.

Figure 6 shows that Net-Pref causes the least number of handovers, as they are mainly initiated when a client moves out of PoA coverage. On the other hand, in the search of the globally optimal allocation, URM and IRM cause a large number of handovers per client, a consequence that is not shared with JURNA, which produces handover rates closer to Net-Pref due to the trigger-based and limited scope of actions.

Additionally, Figure 7 shows that JURNA results in a similar blocking probability to Net-Pref. There is a slightly increased blocking probability for gold and silver users compared to URM, while still below 0.7%. It can be attributed to the proactive network selection in those cases that PoA-overload cannot be immediately triggered, or in the reactive nature of JURNA. On the other hand, Bronze users, with lower traffic demands are more easily accommodated in their pre-selected PoAs, and enjoy lower blocking rates. Since all users share the same resources, it is understandable that blocking a user may result in unblocking one of a different class.

It should be noted that SuperBS serves as the optimal allocation for the fairness index (Figure 4). It is to be evaluated in conjunction with other metrics, as a system that performs equally badly for all clients is perfectly fair. Under this light,

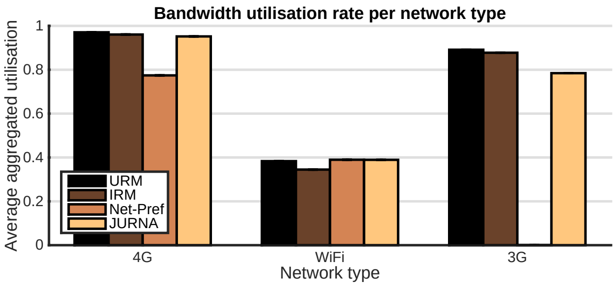

Fig. 5. Average aggregated network utilisation.

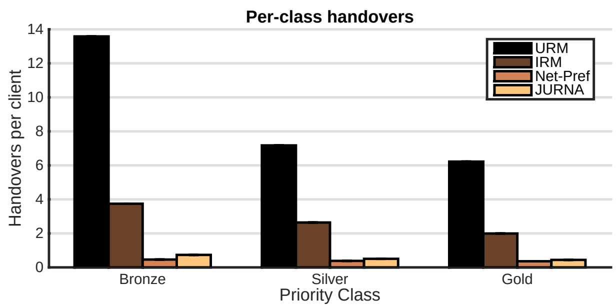

Fig. 6. Average handovers per client.

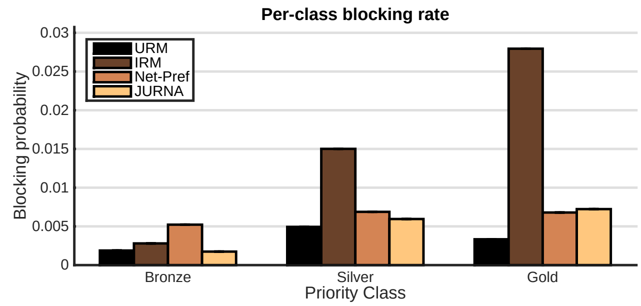

Fig. 7. Average blocking rate (probability of active client blocking).

JURNA performs closely to URM and better than IRM, with significantly fewer handovers. This performance is even more impressive considering JURNA runs in realtime with $O(nS)$ complexity, against URM worst case $O(nS^n)$ complexity and runtime exponential to the number of PoAs (S) and clients (n).

VI. Conclusion - Future Work

Optimising the operation of heterogeneous mobile networks can be attained by using a global view of the network. Still, the high optimisation complexity deems such a centralised solution impractical. JURNA is proposed as a novel utility-based centralised resource and network assignment framework for heterogeneous mobile networks. JURNA employs a novel utility function that considers user and system design goals.

Additionally, a trigger-based scheme is developed to reduce the JURNA operation complexity. JURNA strikes a balance between different design goals including efficient resource utilisation, fair resource allocation, and improved QoS. As future work, we consider further investigation of the scalability of JURNA in very large-scale networks. We also consider integrating new attributes in the design of the utility function. Finally, we plan to work on a detailed system model that allows JURNA to be implemented on actual network equipment.

Acknowledgment

This work is supported by Science Foundation Ireland (SFI) under Research Grant 10RFP/CMS2952 and CTVR under SFI CSET 10|CE|I1853. Portions of the research in this paper used the MDC Database made available by Idiap Research Institute, Switzerland and owned by Nokia.

References

[1] M. Kassar, B. Kervella, and G. Pujolle, "An overview of vertical handover decision strategies in heterogeneous wireless networks," *Computer Communications*, vol. 31, no. 10, pp. 2607–2620, Jun. 2008.

[2] K. Piamrat, A. Ksentini, J.-M. Bonnin, and C. Viho, "Radio resource management in emerging heterogeneous wireless networks," *Computer Communications*, vol. 34, no. 9, pp. 1066–1076, Jun. 2011.

[3] L. Chen, B. Wang, L. Chen, X. Zhang, and D. Yang, "Utility-based resource allocation for mixed traffic in wireless networks," in *Proc. IEEE INFOCOM International Workshop on Future Media Networks and IP-based TV (FMN-IPTV'11)*, Shanghai, China, Apr. 2011, pp. 91–96.

[4] X. Yang, J. Bigham, and L. Cuthbert, "Resource management for service providers in heterogeneous wireless networks," in *Proc. IEEE Wireless Communications and Networking Conference (WCNC'05)*, New Orleans, LA, USA, Mar. 2005, pp. 1305–1310.

[5] I. Tsompanidis, A. H. Zahran, and C. J. Sreenan, "SuperBS: A methodology for resource management in heterogeneous wireless networks," in *Proc. IEEE International Wireless Communications and Mobile Computing Conference (IWCMC'13)*, Cagliari, Italy, Jul. 2013, pp. 1290–1295.

[6] F. Kelly, "Charging and rate control for elastic traffic," *European Transactions on Telecommunications*, vol. 8, no. 1, pp. 33–37, Jan. 1997.

[7] L. Wang and G.-S. Kuo, "Mathematical modeling for network selection in heterogeneous wireless networks - a tutorial," *IEEE Commun. Surveys Tuts.*, vol. 15, no. 1, pp. 271–292, Jan. 2013.

[8] Q.-T. Nguyen-Vuong, Y. Ghamri-Doudane, and N. Agoulmine, "On utility models for access network selection in wireless heterogeneous networks," in *Proc. IEEE Network Operations and Management Symposium (NOMS'08)*, Salvator, BA, Brazil, Apr. 2008, pp. 144–151.

[9] W.-H. Kuo and W. Liao, "Utility-based resource allocation in wireless networks," *IEEE Trans. Wireless Commun.*, vol. 6, no. 10, pp. 3600–3606, Oct. 2007.

[10] A. Abdel-Hadi, M. Ghorbanzadeh, and C. Clancy, "Optimal radio resource allocation for hybrid traffic in cellular networks: Centralized and distributed architecture," *CoRR*, vol. abs/1411.4011, 2014.

[11] V. Srivastava and F. Bullo, "Knapsack problems with sigmoid utilities: Approximation algorithms via hybrid optimization," *European Journal of Operational Research*, vol. 236, no. 2, pp. 488–498, Jul. 2014.

[12] R. Chai, X. Wang, Q. Chen, and T. Svensson, "Utility-based bandwidth allocation algorithm for heterogeneous wireless networks," *SCIENCE CHINA Information Sciences*, vol. 56, no. 2, pp. 95–107, 2013.

[13] I. Tsompanidis, A. H. Zahran, and C. J. Sreenan, "Towards utility-based resource management in heterogeneous wireless networks," in *Proc. ACM International Workshop on Mobility in the Evolving Internet Architecture, (MobiArch'12)*, Istanbul, Turkey, Aug. 2012, pp. 23–28.

[14] N. Kiukkonen, J. Blom, O. Dousse, and J. Laurila, "Towards rich mobile phone datasets: Lausanne data collection campaign," in *Proc. IEEE Int. Conf. on Pervasive Services (ICPS'10)*, Berlin, Germany, Jul. 2010.

[15] J. K. Laurila, D. Gatica-Perez, I. Aad, J. Blom, O. Bornet, T. Do, O. Dousse, J. Eberle, and M. Miettinen, "The mobile data challenge: Big data for mobile computing research," in *Proc. Mobile Data Challenge by Nokia Workshop*, Newcastle, UK, Jun. 2012.

[16] I. Rhee, M. Shin, S. Hong, K. Lee, S. J. Kim, and S. Chong, "On the levy-walk nature of human mobility," *IEEE/ACM Trans. Netw.*, vol. 19, no. 3, pp. 630–643, Jun. 2011.

[17] I. Tsompanidis, A. Zahran, and C. Sreenan, "Mobile network traffic: A user behaviour model," in *Proc. IFIP Wireless and Mobile Networking Conference (WMNC'14)*, Vilamoura, Portugal, May 2014, pp. 1–8.