

Ensemble Techniques for Scheduling in Heterogeneous Wireless Communications Networks

David Lynch, Michael Fenton, Stepan Kucera, Holger Claussen, and Michael O'Neill

Abstract Operators deploy Small Cells in high traffic regions to boost the capacity of their wireless networks. However, User Equipments (UEs) at Small Cell edges experience severe interference from neighbouring high-powered Macro Cells. A fair trade-off between cell-edge and cell-centre performance can be realised by intelligently scheduling Small Cell attached UEs to receive data. Grammar-based Genetic Programming is employed to learn models which map measurement reports to schedules on a millisecond timescale. The evolved models are instrumented as ensembles. The proposed system significantly outperforms a state of the art benchmark algorithm and is within 7.5% of the estimated optimum.

1 Introduction

Traditional single-tiered cellular networks are struggling to cope with exponentially rising demand [1]. Capacity can be increased by supplementing the existing Macro Cell (MC) tier with lower-powered Small Cells (SCs) [2]. SCs provide a local capacity boost in traffic hotspots. The resulting two-tiered configuration is known as a Heterogeneous Network or 'HetNet'.

HetNets are spectrally efficient because both cell tiers reuse the same scarce and expensive bandwidth. Unfortunately, co-channel operation results in significant interference at the edges (i.e. borders) of SCs. Cell-edge interference is alleviated by periodically muting MCs. A MC broadcasts only minimal reference signals in these co-called 'Almost Blank Subframes' (ABSF) [5]. Note that a 'subframe' is a 1 ms interval in which cells send packets to their attached User Equipments (UEs: smartphones, tablets etc.). Interference is drastically reduced for SC attached UEs when interfering MCs execute an ABSF.

Operators (e.g. Vodafone Group plc) must ensure that UEs do not experience significant packet losses, which will result in poor customer satisfaction and churn. Network resources can be more fairly distributed by scheduling cell-centre UEs when MC interference is higher (e.g. in non-ABSFs), so that quieter airtime (e.g. ABSFs) is reserved for cell-edge UEs. Hence, high-performing cell-centre UEs can be sacrificed for the sake of cell-edge UEs.

However, optimally allocating UEs between ABSFs and non-ABSFs is a non-trivial task. Furthermore, schedules must be inferred from measurement reports on a millisecond timescale. Previous work by the authors [6] has demonstrated the suitability of Grammar-based Genetic Programming (GBGP) [7] as a framework for automatically evolving SC schedulers. This paper proposes an ensemble technique which exploits the stochasticity of evolutionary search to realise a further 9.4% performance dividend.

D. Lynch (e-mail: david.lynch.1@ucdconnect.ie) · M. Fenton · M. O'Neill
Natural Computing Research and Applications Group, School of Business, University College Dublin

S. Kucera · H. Claussen
Bell Laboratories, Nokia-Ireland

2 Problem Definition

UEs send reports of $Q_{u,f}^t$ to their serving cell after every ‘frame’ consisting of $|\mathcal{F}| = 8$ subframes¹. Shannon’s formula [8] gives the rate at which information flows through a wireless channel to UE u , in subframe f of frame t :

$$R_{u,f}^t = \frac{B}{N_f^t} \times \log_2(1 + SINR_{u,f}^t), \quad (1)$$

where, $R_{u,f}^t$ is the downlink rate, $B = 20$ MHz is the available bandwidth, N_f^t is the number of UEs receiving data from u ’s serving cell in f and $SINR_{u,f}^t$ is the signal u receives from its serving cell in f divided by the interference from all other cells plus background noise. Let $Q_{u,f}^t := \log_2(1 + SINR_{u,f}^t)$ denote the channel quality experienced by u in subframe f of frame t .

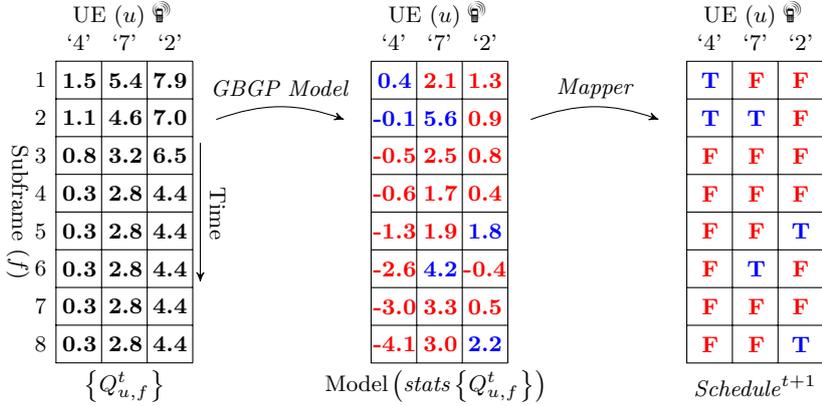


Fig. 1: Generating schedules from measurement reports.

The leftmost panel of Fig. 1 displays typical values of $Q_{u,f}^t$ over frame t , for a SC (s) with three attached UEs (let \mathcal{A}_s^t denote the set of UEs attached to s in frame t). UEs and subframes are represented by columns and rows respectively. GBGP is employed to learn a mapping from statistics over the set $\{Q_{u,f}^t | u \in \mathcal{A}_s^t, f \in \mathcal{F}\}$ to a schedule for s . The real-valued outputs of the model (central panel) are interpreted as a Boolean schedule (rightmost panel), which s will observe in frame $t+1$. Each UE is forced to receive data in exactly two subframes by setting the largest two cells in each column from the central panel to ‘True’ and the remaining cells to ‘False’ [6]. For instance, UE4 will receive packets from s in subframes $f = \{1, 2\}$ but not in $f = \{3..8\}$.

The quality or “fitness” of the schedule in Fig. 1 is given by the sum-log-rates (SLR) metric [3]:

$$SLR_s := \sum_{u \in \mathcal{A}_s^{t+1}} \log_e \left(\frac{1}{|\mathcal{F}|} \sum_{f=1}^{|\mathcal{F}|} R_{u,f}^{t+1} \right). \quad (2)$$

Eq. 1 implies that $R_{u,f}^t \propto Q_{u,f}^t / N_f^t$. Therefore, knowledge of the reported channel qualities and schedule are sufficient to evaluate Eq. 2. For example, $SLR_s = 48.72$ where s is the SC depicted in Fig. 1 (assuming for the sake of clarity that $Q_{u,f}^{t+1} = Q_{u,f}^t$ and $\mathcal{A}_s^{t+1} = \mathcal{A}_s^t$). A scheduler is optimal if it maximises SLR_s for all SCs s in the network in an arbitrary frame.

¹ In fact, $|\mathcal{F}| = 40$ but it is sufficient to compute schedules for $f = \{1..8\}$ only, as explained in [6].

3 Previous Work

This section reviews a state of the art benchmark algorithm for scheduling in Het-Nets proposed by López-Pérez and Claussen (2013) [4]. See [6] for a more complete review of the literature.

The benchmark initialises two queues, $\mathcal{Q}_{non-ABS F}$ and $\mathcal{Q}_{ABS F}$, for each SC s . UEs in the former queue are only scheduled during subframes in which the nearest MC transmits; those in the latter queue receive data only during subframes in which the nearest MC mutes. Initially, the $\lfloor \mathcal{A}_s^t/2 \rfloor$ UEs with highest average channel quality are placed in $\mathcal{Q}_{non-ABS F}$, and the remaining worst performers occupy $\mathcal{Q}_{ABS F}$. Eq. 1 is evaluated for all UEs in both queues and their average downlink rates are computed. The worst UE (w.r.t average rate) is identified in each queue. UEs are moved between queues, subject to constraints, in order to equalise the rates of the worst performers. Once the desired queue lengths stabilise, the SC schedules $u \in \mathcal{Q}_{non-ABS F}$ in non-ABSFs and $u \in \mathcal{Q}_{ABS F}$ in ABSFs. The benchmark thus tends to sacrifice the best performing UEs (by scheduling them in non-ABSFs) so that cell-edge UEs profit from the reduced interference in ABSFs.

4 Experiments

An 3.61 km² area of downtown Dublin was simulated by modelling the distribution of buildings, open spaces and waterways to compute signal propagation path losses. Training data were generated from a network with 30 SCs and 21 MCs. A single training case was the set $\{Q_{u,f}^t | u \in \mathcal{A}_s\}$ for SC s . Three hundred training cases were saved over ten frames in order to evolve models that generalised well. Cell powers and MC muting patterns were set according to the heuristics in [4], thus ensuring a fair benchmarking of the evolved models.

A symbolic regression style grammar, identical to that outlined in [6], was used to evolve functional expressions with GBGP. The same evolutionary parameters were adopted from [6] except $\#gens := 200$. The best models from 1500 independent runs were arbitrarily grouped into 30 ensembles (each consisting of 50 models).

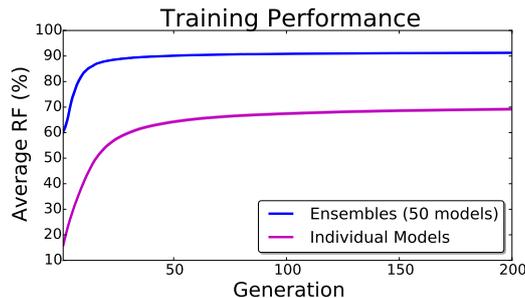


Fig. 2: Performance on training data.

4.1 Training

The performance of a GBGP model on a single SC s is given by,

$$RF_{s,model} := \left(\frac{SLR_{s,model} - SLR_{s,baseline}}{SLR_{s,CMA} - SLR_{s,baseline}} \right) \times 100\%, \quad (3)$$

where, $RF_{s,model}$ expresses the ‘relative fitness’. In Eq. 3, $SLR_{s,model}$ is the SLRs realised by the model-generated schedule. Similarly, $SLR_{s,baseline}$ is the SLRs if UEs are greedily scheduled in every subframe (they are unscheduled if $SINR_{u,f} \leq$

−5.5 dB). The Covariance Matrix Adaptation-Evolutionary Strategy (CMA) is executed once off-line for all SCs in the training set to estimate the optimum schedules. Hence, $SLR_{s,CMA}$ estimates the achievable SLRs for case s . Eq. 3 evaluates to 100% if the model generates an optimal schedule and $\leq 0\%$ if it cannot beat the greedy baseline. Overall fitness is defined as the average of $RF_{s,model}$ across all SCs in the training set. Fig. 2 illustrates how the average relative fitness converges on the training cases over 200 generations. The ensembles significantly outperform individual models on training data throughout the evolutionary runs.

4.2 Performance on Test Data

To account for finite processing time, schedules for frame $t+1$ must be computed in frame t based on $\{Q_{u,f}^t\}$. The 40 ms interval between t and $t+1$ is sufficient to execute an ensemble consisting of multiple models. Each model generates a hypothesis schedule for SC s in t . In real time, Eq. 2 is evaluated for each hypothesis against $\{Q_{u,f}^t|u \in \mathcal{A}_s^t\}$ and the best schedule (w.r.t SLRs) is used by s in $t+1$. The proposed ensemble method exploits the stochastic nature of GBGP since evolved models tend to admit non-overlapping errors. This occurs because they occupy different local optima in the search space defined by the grammar.

The UE mobility model is as follows. In the interval between t and $t+1$, 10% of the 1260 UEs are displaced 0.54 m ($\equiv 50$ km/h) in a random direction, 45% are displaced 0.056 m ($\equiv 1.4$ m/s) and the remaining UEs move 0.01 m. This mobility model simulates the effects of users driving and walking across the map, or else experiencing channel quality fluctuations while static.

Tab. 1 compares the benchmark and evolutionary methods on unseen test data. Eq. 3 was evaluated for all SCs (test cases) over 100 frames and averaged. Comparing Tab. 1 with Fig. 2 we see that the evolved models generalise very well to unseen cases. A one-way ANOVA reveals that there is a significant difference between the methods and Tukey’s post-hoc analysis suggests that each group mean is significantly different from the others. Schedules computed by CMA in t are on average suboptimal by a factor of 1.2% in $t+1$. This makes sense since they are based on slightly outdated reports. The ensemble is within 8.1% of the estimated optimum which is impressive since it executes in under 40 ms, unlike CMA which takes several seconds. Highly fit ‘best-of-ensemble’ models are outperformed by ensembles which can generate tailored schedules for arbitrary cases. Finally, the benchmark achieves much lower SLRs and higher variance than all evolved models.

	Benchmark	Best-of-Ensemble	Ensemble	CMA
RF^{avg} (%)	20.2 ± 8.2	84.0 ± 2.6	91.9 ± 1.4	98.8 ± 2.4

Table 1: Average relative fitness of the methods over 100 test frames.

Fig. 3 (a) plots the average performance of the 30 ensembles on the test set against ensemble size. Only three models working cooperatively are needed to surpass the best individual model from 50 runs. The grey curve describes thirty models that were selected at random. The relatively wide 95% confidence interval that is suggested by the grey shaded region, underscores the need for many independent runs when building a scheduler using GBGP. Execution time increases linearly with ensemble size and is well under 40 ms for $size = 50$. Better results would be achieved with larger ensembles since fitness has not fully converged after 200 generations.

Fig. 3 (b) breaks down the performance of each method on the test set for different SC loads. The estimated optimum is given by running CMA on $\{Q_{u,f}^{t+1}\}$ but, as before, all other methods compute schedules in t based on $\{Q_{u,f}^t\}$. The ensemble and best-of-ensemble models significantly outperform the benchmark across

almost all cell loads. The evolved models are much more stable compared to the benchmark. Nonetheless, a clear optimality gap exists between the ensemble and CMA. This illustrates an opportunity for smarter joint optimisation of the ensemble members in future work, so that they cooperate better as a collective.

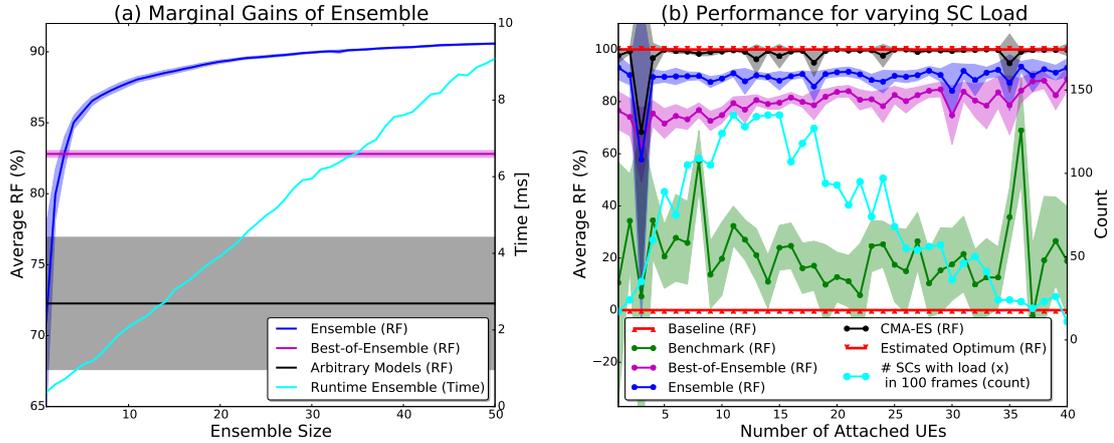


Fig. 3: Performance w.r.t ensemble size (left) and cell load (right).

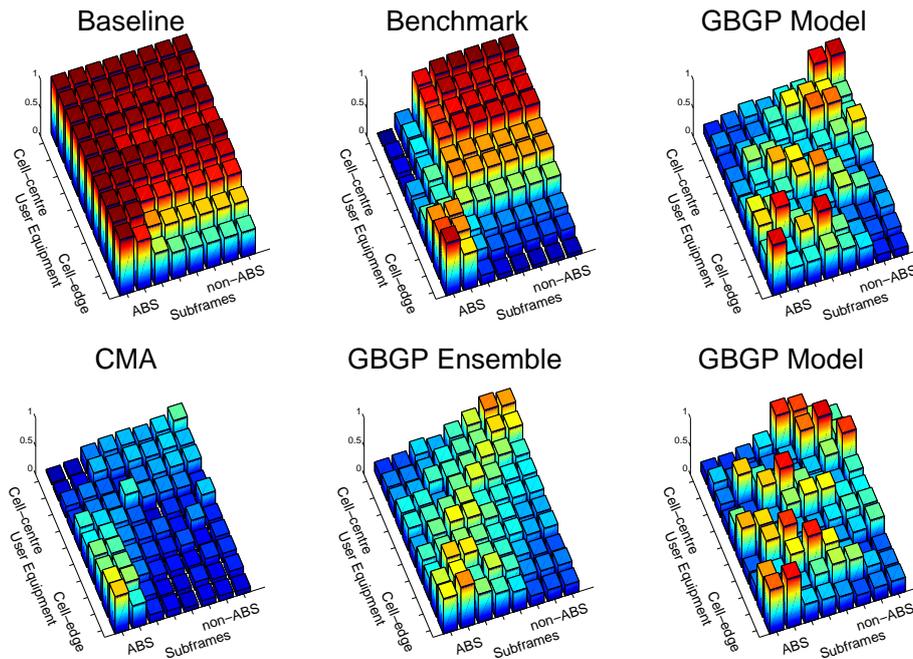


Fig. 4: Visualising the semantics.

5 Discussion

It is instructive to visualise the semantics of each method via heatmaps. Fig. 4 summarises the scheduling decisions made by the various methods for 75 (unseen) SCs with exactly ten attached UEs. Deep red in cell (u, f) indicates that u is scheduled in f for all 75 cases. Conversely, deep blue implies that a method never schedules u in f . UEs and subframes are represented by columns and rows respectively. UEs are sorted with respect to average channel quality from low (edge) to high (centre).

The greedy baseline schedules each UE in all subframes. Hence, its heatmap is mostly deep red with some lighter colours corresponding to cases where UEs that cannot be scheduled because $SINR_{u,f} \leq -5.5$ dB. The benchmark displays a more

intelligent strategy. Cell-centre UEs are denied data in the ABS subframes thus liberating bandwidth for cell-edge UEs when the channel quality is high. The converse occurs in subframes 3...8 when more MCs are active. Therefore, the benchmark reserves premium airtime for highly interfered UEs at the expense of top performers. The latter are compensated with a larger number of low-channel quality subframes in which edge UEs are unscheduled.

The rightmost heatmaps were generated by arbitrarily selected GSGP models. The benchmark's core strategy is discernible but the peculiar semantics of each GSGP model are quite different. Consequently, hypotheses generated by different models tend to be dispersed in the space of possible schedules for a SC. Thus, the ensembles outperform individual highly fit generalising models. Finally, the small optimality gap between the ensembles and CMA is echoed by their similar semantics. The light blue palette characterising both heatmaps suggests that these methods fit highly specialised schedules. This degree of specialisation is unattainable for a single generalising model.

6 Conclusions and Future Work

It is feasible to evaluate more than one model to generate several hypothesis schedules for a SC in the intervals between successive frames. The ensemble members must make non-overlapping errors for this approach to work well. GBGP tends to yield semantically diverse solutions so that the hypotheses are well dispersed in the space of possible schedules. This allows an ensemble of GBGP models to closely approximate the highly optimised schedules given by CMA. Crucially, the proposed solution executes on a short enough timescale so that new schedules can be generated in every frame.

Schedulers which cooperate effectively emerge naturally from the evolutionary search process. Future work could explicitly encourage cooperation among evolving ensemble members through the fitness function. This could be enabled by re-weighting difficult cases in a manner analogous to the machine learning technique of boosting.

Acknowledgements This research is based upon works supported by the Science Foundation Ireland under grant 13/IA/1850.

References

1. Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2015-2020. Cisco, 2016.
2. Alcatel-Lucent. Small Cell Solutions. <https://www.alcatel-lucent.com/solutions/small-cells>, 2015. [Report].
3. S. Deb, P. Monogioudis, J. Miernik, and J. P. Seymour. Algorithms for enhanced inter-cell interference coordination (eICIC) in LTE HetNets. *IEEE/ACM Transactions on Networking (TON)*, 22(1):137–150, 2014.
4. D. López-Pérez and H. Claussen. Duty Cycles and Load Balancing in HetNets with eICIC Almost Blank Subframes. In *Personal, Indoor and Mobile Radio Communications (PIMRC Workshops), 2013 IEEE 24th International Symposium on*, pages 173–178. IEEE, 2013.
5. D. Lopez-Perez, I. Guvenc, G. De la Roche, M. Kountouris, T. Q. Quek, and J. Zhang. Enhanced intercell interference coordination challenges in heterogeneous networks. *IEEE Wireless Communications*, 18(3):22–30, 2011.
6. D. Lynch, M. Fenton, S. Kucera, H. Claussen, and M. O'Neill. Evolutionary learning of scheduling heuristics for heterogeneous wireless communications networks. In *Proceedings of the 2016 on Genetic and Evolutionary Computation Conference*, pages 949–956. ACM, 2016.
7. R. I. McKay, N. X. Hoai, P. A. Whigham, Y. Shan, and M. O'Neill. Grammar-based Genetic Programming: a survey. *Genetic Programming and Evolvable Machines*, 11(3-4):365–396, 2010.
8. C. E. Shannon. Communication In The Presence Of Noise. *Proceedings of the IRE*, 37(1):10–21, 1949.